University of Alberta

Analysis of the mechanism of action of TraJ in regulating F plasmid gene expression

by

J. Manuel Rodriguez-Maillard ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

in

Microbiology and Biotechnology

Department of Biological Sciences

Edmonton, Alberta

Fall, 2008

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-47401-3
Our file Notre référence
ISBN: 978-0-494-47401-3

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

The F plasmid encodes a type four secretion system involved in conjugation. F TraJ is a positive activator of the major transfer operon, relieving H-NS repression. Analysis of F and F-like TraJ indicated clusters of conserved amino acid residues including a predicted DNA-binding helix-turn-helix motif. Mutations of conserved amino acids in the helix-turn-helix region diminish conjugation. Moreover, a DNA band corresponding to the *traYX* promoter was detected from a chromatin immunoprecipitation assay, only in a strain producing TraJ. Therefore, TraJ is a DNA binding protein that possibly counteracts H-NS by competing for binding sites. TraJ C-terminus is crucial for conjugation as deletion of the last four residues disrupts conjugation. TraJ forms dimers and is 226, not 229 amino acids long as previously considered. TraJ seems to be part of the MarR/SlyA family of regulators, due to several similarities with its member RovA.

Acknowledgements

Many people have made a difference in my life and I would like to thank each and everyone who helped me achieve my degree all the way up here in Canada.

I would like to thank my supervisor Laura S. Frost for giving me a place to work and pursue my degree.

I can never thank my family enough for building me up as a person and their never ending support all throughout my studies.

I would also like to thank Jan, Barbara, Isabella C.Y., Ryan, Andrew N., Bernie, Denis and other members of the frost lab for all the assistance whether helping with experiments or in answering all sorts of questions.

Finally I would like to thank my committee members for taking the time to appraise my work.

Table of Contents

Chapter 1. Introduction.....1

A. The F plasmid.....2

I. Proteins encoded in the F transfer region.....5

II. Conjugation of the F plasmid.....6

III. F-like plasmids.....9

B. Regulation of F transfer.....10

I. TraJ, the positive regulator.....10

II. The Fertility Inhibition (FinOP) system and IS3.....12

III. The Histone-like nucleoid structuring protein (H-NS).....13

IV. Other regulators of F.....14

C. Gene silencing.....18

I. Mechanisms of H-NS desilencing.....22

II. Nucleoid-associated proteins.....24

D. DNA binding proteins.....25

I. Ribbon-Helix-Helix (RHH).....26

II. Helix-Turn-Helix (HTH).....27

III. Winged-Helix domain.....28

E. Research objectives.....29

Chapter 2. Materials and Methods.....30

A. Strains and plasmids.....31

B. Chemicals, growth media and enzymes.....43

C. Site-directed mutagenesis.....43

D. Deletion mutagenesis.....44

E. Mating assays.....45

F. Immunoblot analysis.....46

G. Chromatin Immunoprecipitation (ChIP) assay.....46

H. Mass spectrometry.....48

I. Cross-linking with DSS.....49

Chapter 3. TraJ is a DNA binding protein with an essential four amino acid C-terminus tail.....50

A. Introduction.....51

B. Results.....52

I. Sequence alignment of TraJ from F-like plasmids and determination of a putative DNA-binding site.....52

II. Site-directed mutagenesis of specific residues in the putative HTH site in F TraJ.....57

III. F TraJ C-terminal deletion mutagenesis.....61

IV. Wild-type F TraJ *in vivo* DNA binding assays.....65

C. Discussion.....70

Chapter 4. TraJ is a 226 amino acid protein that forms dimers and higher-order complexes.....76

A. Introduction.....77

B. Results.....78

I. Mass Spectrometry data for F TraJ.....78

II. F TraJ translation start studies.....81

III. TraJ oligomerization in an F^+ and F^- strain.....84

C. Discussion.....89

Chapter 5. Discussion and future directions.....94

A. Discussion.....95

B. Future directions.....104

Chapter 6. References.....108

List of figures

Figure 1.1 Map for the F plasmid (Genbank accession: AP001918).....4
Figure 1.2 F plasmid conjugative transfer.....8
Figure 1.3 The effect of H-NS and the three *tra* regulatory proteins on the main promoters of the F transfer region.....16
Figure 1.4 Silencing of a promoter region by H-NS.....21
Figure 1.5 Sequence comparison between the inverted repeat of the *traJ-traY* intergenic region of R100 and F.....29
Figure 3.1 Sequence alignment of F TraJ (Accession: BAA97942) and two orthologues from plasmids R100 (Accession: BAA78849), R1-19 (Accession: AAA92657).....54
Figure 3.2 Sequence alignment of F TraJ (Accession: BAA97942) and seven homologues from different strains and plasmids (Accession: ZP_03035127, NP_490561, AAA25523, ZP_03031175, YP_001919411, YP_001711968, YP_001338611).....56
Figure 3.3 Immunoblots of the samples used in the mating assays for the TraJ point mutants expressed from pBAD24.....60
Figure 3.4 Immunoblots of the samples used in the mating assays for the TraJ deletion mutants expressed from pBAD24.....64
Figure 3.5 Sequence of F *tra* DNA involved in the ChIP assay.....67
Figure 3.6 PCR results in duplicate from the ChIP assay of TraJ run on a 1.5% agarose gel.....69
Figure 3.7 Sequence comparison between the C-terminal regions of F TraJ (Accession: BAA97942), RovA (Accession: CAH21526) and SlyA (Accession: AAL20366).....74
Figure 4.1 Mass spectrometry results for the TraJ sample obtained from the strain MC4100/F*lac traJ90*/pBAD24TraJ at three hours of growth.....80

Figure 4.2 Immunoblots of the TraJ methionine mutants expressed from pBCSK$^+$ with controls.....83
Figure 4.3 Cross-linking with DSS for MC4100/F*lac traJ90* samples with either wild-type TraJ (A) or truncated (Δ6 or Δ30) versions of TraJ (B).....86
Figure 4.4 Cross-linking with DSS for PD32 (*hns*$^-$) samples with either wild-type TraJ from F*lac* (A) or TraJ (wild-type or Δ30) expressed from the pBAD33 vector with no F present (B).....88
Figure 4.5 Model for the TraJ oligomeric state in both exponential and stationary phase.....92
Figure 5.1 Summary of most of the mutants used in this study.....97
Figure 5.2 Model for TraJ regulation mechanism.....99
Figure 5.3 Model for silencing-desilencing of F transfer region.....103

List of Tables

Table 2.1 List of *Escherichia coli* K-12 strains used.....35
Table 2.2 List of plasmids used.....36
Table 2.3 List of oligonucleotides used.....38
Table 3.1 Mating assay results for the TraJ point mutants in the putative DNA-binding domain.....58
Table 3.2 Mating assay results for the TraJ C-terminus deletion mutants.....62

Abbreviations

α	Alpha
A	Adenine
Ap	Ampicillin
ATP	Adenosine triphosphate
β	Beta
bp	Base pair
C	Cytosine
cAMP	Cyclic adenosine monophosphate
Cbp	Curved DNA-Binding Protein
ChIP	Chromatin immuno-precipitation
Cm	Chloramphenicol
Cpx	Conjugative Pilus Expression
CRP	cAMP receptor protein
Δ	Delta (deletion)
ddH_2O	Double distilled water
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleoside triphosphate
Dps	DNA-binding Protein from Starved cells
DRE	Downstream regulatory element
DSS	Disuccinimidyl suberate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
EPEC	Enteropathogenic *E.coli*
ESI-TOF	Electrospray ionization time-of-flight
F	Fertility sex factor
F^+	F conjugative strain

F⁻	F non-conjugative strain
Fin	Fertility inhibition
Fis	Factor of inversion stimulation
G	Guanine
h	Hours
Hfq	Host factor for Qβ replication
Hha/Ymo	Haemolysin/*Yersinia* modulator
H-NS	Histone-like nucleoid structuring
H-NST	Truncated H-NS
HTH	Helix-turn-helix
HU	Heat unstable
Ici	Inhibitor of Chromosome Initiation
IHF	Integration host factor
Inc	Incompatibility
inv	Invasin
IS	Insertion sequence
Km	Kanamycin
Kb	Kilobase
kDa	Kilodalton
Lac	Lactose
LB	Luria-Bertani
Lrp	Leucine responsive protein
M	Molar
mg	Milligram
µg	Microgram
ml	Milliliter
µl	Microliter
mM	Millimolar
µmol	Micromolar

mRNA	Messenger RNA
MS	Mass spectrometry
nic	Nick site
nM	Nanomolar
OD_{600}	Optical density at 600 nm
Omp	Outer membrane protein
orf	Open reading frame
oriS	Origin of replication S
oriT	Origin of transfer
oriV	Origin of replication V
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
P_{finP}	*finP* promoter
pmol	picomole
P_J	*traJ* promoter
P_M	*traM* promoter
P_{traS}	*traS* promoter
P_{trbF}	*trbF* promoter
P_Y	*traY* promoter
R	Resistance factor
RBS	Ribosome binding site
Rep	Replication
RHH	Ribbon-helix-helix
RNA	Ribonucleic acid
RNAP	RNA polymerase
RNase	Ribonuclease
Rov	Regulator of virulence
sbj	Site of binding for TraJ
sbm	Site of binding for TraM

sby	Site of binding for TraY
SDS	Sodium dodecyl sulphate
Sfh	*Shigella flexneri* H-NS-like
Sfr	Sex factor regulator
SL	Stem loop
Sp	Spectinomycin
ssDNA	Single-stranded DNA
Stp	Suppressor of the *td*$^{-}$ Phenotype
SUMO	Small Ubiquitin-like Modifier
T	Thymine
T4SS	Type four secretion system
TEMED	Tetramethyl-ethylenediamine
Tn	Transposon
TOPO	Topoisomerase
Tra	Transfer
Tris	Tris (hydroxymethyl) aminomethane
U	Uracil
UPEC	Uropathogenic *E.coli*
URE	Upstream regulatory element
UV	Ultraviolet
Vir	Virulence
Ydg	YmoA/Hha paralogue
°C	Degrees Celsius

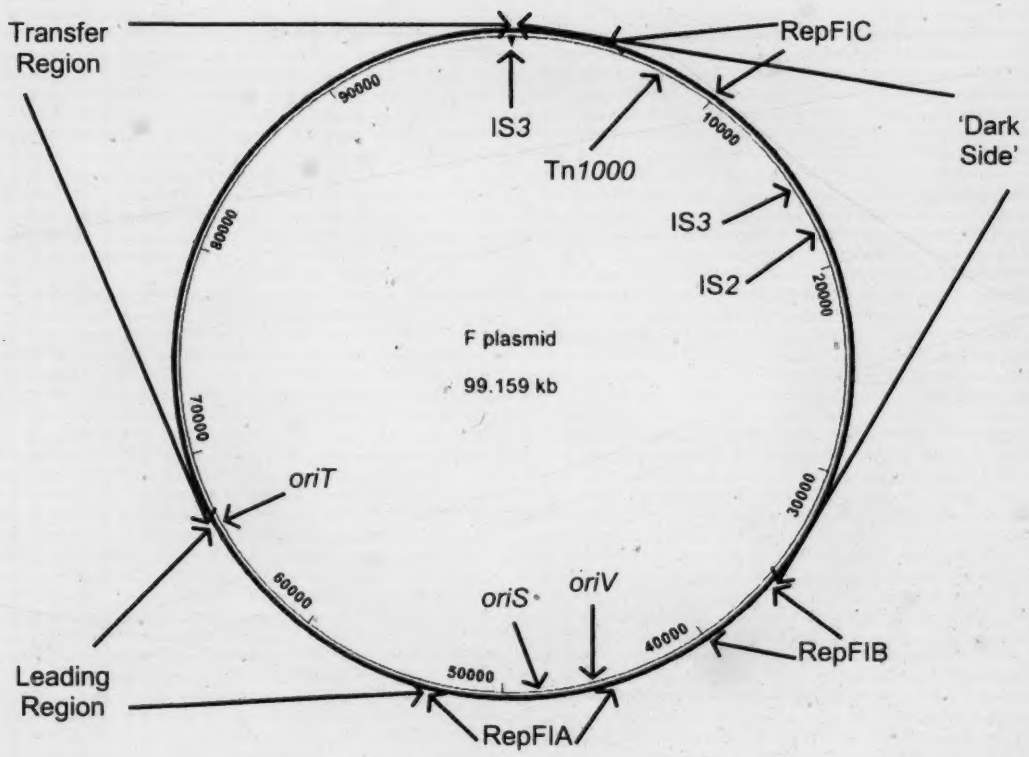
Chapter 1. Introduction

A. The F plasmid

F, also known as the Fertility factor, is a 100 kb plasmid that was demonstrated to mediate conjugation, which was first described in *Escherichia coli* K-12 (Tatum & Lederberg, 1947). F has been thoroughly studied since then but important findings are still being recorded. There are marked differences between conjugation of different plasmids, which are categorized into different Incompatibility (Inc) groups. This incompatibility phenomenon was originally observed over 40 years ago for the Resistance (R) plasmids (Watanabe *et al.*, 1964). The F plasmid is part of the IncF family, which is also subdivided into seven different groups (FI-FVII, collectively know as F-like) based on their replicons. As shown in figure 1.1, the F plasmid includes a 15 kb leading region, a 'dark side' mostly composed of genes of unknown function, three (two complete and one incomplete) Rep regions for autonomous replication and four transposable elements which include Tn*1000*, insertion sequence *2* (IS2) and two copies of an IS*3* (*a, b*) (Manwaring *et al.*, 1999). Finally, it includes a 33.3 kb transfer region, which encodes all the elements required for conjugation (Frost *et al.*, 1994).

The F plasmid encodes all the factors required for a true type-four secretion system (T4SS), a complex in the cell membrane which is able to transfer DNA as well as proteins to the recipient cell (Lawley *et al.*, 2003). Moreover, F is not only able to transfer itself, the same T4SS machinery can be used by other mobilizable elements residing in the donor cell that are incapable of conjugation themselves (Reeves & Willetts, 1974). Both F and these mobilizable elements are also capable of carrying chromosomal fragments or transposons. In addition, recombination between F and these other elements has also been shown to occur at a faster rate than between

Figure 1.1 Map for the F plasmid (Genbank accession: AP001918). Total size is indicated in the middle of the map, while numbers around it indicate base pair coordinates starting from the small triangle on top of the map. Main regions (transfer (Tra), leading, 'dark' and replication (Rep) regions) are enclosed by arrows outside of the map. Insertion sequences, transposons and origins of transfer and replication are indicated by arrows inside the map.


Transfer Region
RepFIC
'Dark Side'
IS3
Tn1000
IS3
IS2
F plasmid
99.159 kb
oriT
oriS
oriV
RepFIB
Leading Region
RepFIA
10000
20000
30000
40000
50000
60000
70000
80000
90000

F and the chromosome (Boyd *et al.*, 1996), thus promoting adaptation and driving evolution in bacteria. Conjugation is an energetically expensive process since it requires ATP hydrolysis in a number of steps (Majdalani & Ippen-Ihler, 1996); therefore, F transfer is tightly regulated by a plethora of host- and plasmid-encoded factors. This regulation is especially useful in times of nutritional and environmental stress, when the cells need to conserve metabolic energy. Unlike most of the F-like plasmids, F has a derepressed phenotype due to the fertility inhibition gene *finO*, which is interrupted by the IS*3a* element (Cheah & Skurray, 1986). Nevertheless, several other factors still manage to regulate F transfer during stress conditions or in stationary phase.

I. Proteins encoded in the F transfer region

The F transfer region is a 33.3 kb segment of the F plasmid which encodes proteins for pilus synthesis and assembly, nicking and transport, mating pair stabilization, surface exclusion and regulation. It also includes genes of yet unknown function and an IS3 sequence (Frost *et al.*, 1994). Pilus biogenesis and maintenance include the genes: *traA* (pilin), *traL, traE, traK, traB, traP, traV, traC, trbI, traW, traU, trbC, traF, traQ, trbB, traH, traG* (N-terminal region) and *traX*. The nicking and transport group include: *oriT, traM, traY, traD* and *traI*. Mating pair stabilization genes include *traU, traN* and the C-terminal region of *traG* (Firth & Skurray, 1992; Moore *et al.*, 1990). Entry and surface exclusion genes include *traS* and *traT* (Jalajakumari *et al.*, 1987). Regulation genes include *traM, traJ, finP, traY* and *finO* (Frost *et al.*, 1994). Genes with unknown function include: *artA, traR, trbA, trbD, trbE, trbF, trbG, trbH* and *trbJ* (Frost *et al.*, 1994). Plasmids from IncF share homology within their transfer regions; however, in regard to regulation and DNA metabolism, plasmid-specific genes are found (Anthony *et al.*, 1999).

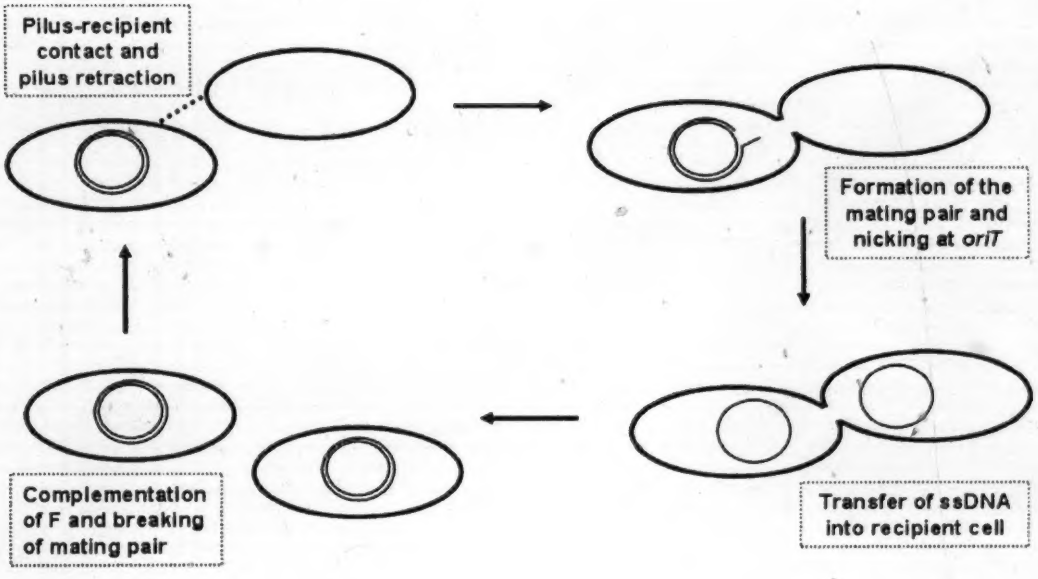
Both *traM* and *traJ* have their own promoters, P_M ($_1$ and $_2$) and P_J respectively (Penfold *et al.*, 1996). *finP* and *artA*, the antisense genes, also have their own promoter regions (van Biesen & Frost, 1994; Wu & Ippen-Ihler, 1989). All the remaining transfer genes are part of a long (~33 kb) operon spanning from *traY* to *traX* with P_Y as the main promoter. Secondary promoters can also be found for *trbF, traI, traS, traT* and *traD*, but their impact is limited (Ham *et al.*, 1989). Two of these secondary promoters, P_{trbF} and P_{traS}, are located within the coding region of their respective upstream gene, *trbJ* and *traG*.

II. Conjugation of the F plasmid

Transfer starts when the tip of the F pilus finds an F^- recipient cell. This contact causes pilus retraction and the two cells are brought into close proximity. The specific mechanism involved in depolymerization of the pilin subunits is yet undefined, as is the final destination of the TraA monomers disassembled from the pilus. Once cells come into contact with each other, a channel is opened between them, this process is called the formation of the mating pair (figure 1.2). The F mating pair involves the proteins TraN and TraG in the outer and inner membranes respectively, as well as OmpA in the recipient (Klimke & Frost, 1998). TraU is also considered to play a role in the stability of the mating pair, as *traU* mutants show a phenotype similar to the *traG* and *traN* mutants (Moore *et al.*, 1990). Entry and surface exclusion proteins, TraS and TraT, can prevent multiple simultaneous matings that might lead to membrane damage (Achtman *et al.*, 1977).

Nicking of the double-stranded DNA occurs continuously at the F plasmid *nic* site, found within the origin of transfer (*oriT*). Nicking is

Figure 1.2 F plasmid conjugative transfer. An F^+ cell produces a pilus that comes into contact with an F^- recipient cell. The pilus retracts and brings the cells together, forming the mating pair. F DNA is then nicked at the *nic* site in *oriT* and single-stranded DNA (ssDNA) is transferred to the recipient in a 5' to 3' direction with the relaxase bound to the 5' end (Matson & Ragonese, 2005). The F ssDNA is replicated and the mating pair is ruptured. Finally both cells are now F^+ and are ready for further rounds of conjugative transfer.


Pilus-recipient contact and pilus retraction
Formation of the mating pair and nicking at oriT
Transfer of ssDNA into recipient cell
Complementation of F and breaking of mating pair

achieved by the relaxosome, which includes the TraY, TraM and Tral (relaxase) proteins as well as the integration host factor (IHF) bound to the DNA at *oriT* (Howard *et al.*, 1995). However, once the mating pair is formed, one strand of the nicked plasmid is transferred by a mechanism that resembles rolling circle replication into the recipient cell. DNA is transferred from the 5' to 3' end and takes approximately five minutes under optimal conditions (Frost *et al.*, 1994). The transferred single-stranded DNA (ssDNA) becomes double-stranded once its complement is synthesized inside the new host, so it can be ready to further transfer itself. Several elements from the F leading region are essential for the establishment of the ssDNA, therefore they are transferred first into the recipient. These elements include: F*rpo* which is a site for ssDNA replication initiation (Masai & Arai, 1997), *flmABC* which is a system involved in maintaining unstable plasmids (Manwaring *et al.*, 1999), *psiAB* which is a protein system that inhibits SOS response in the recipient (Dutreix *et al.*, 1988); (Loh *et al.*, 1990), and *ssb* (*ssf*) which encodes a ssDNA binding protein that promotes replication of the plasmid (Chase *et al.*, 1983).

III. F-like plasmids

Transmissible plasmids are grouped into incompatibility (Inc) groups due to the inability of different plasmids to coexist in a single cell. F belongs to the IncF1 subgroup; other well-studied groups include IncP, IncC, IncN, IncW, IncQ, IncD, IncJ and IncI. Another classification includes only two groups based solely on pilus length and flexibility (short-rigid and long-flexible) as well as its sensitivity to phages (Bradley, 1980). In this alternative classification, the F pilus is considered to be the long-flexible kind. F is the first F-like plasmid with its sequence completely determined (Genbank accession: AP001918). Examples of plasmids from the other

subdivisions include R100, pSU316, R124 from the Inc families FII, III and IV respectively (Genbank accession: AP000342, M26937, K03092). F-like plasmids generally encode important elements such as antibiotic resistance, toxins or additional metabolic capabilities.

Most genes in the transfer region of the F-like plasmids share homology among the structural proteins (Lawley *et al.*, 2003). However, nicking and regulation genes are not generally conserved, as is the case of *traI, traM, traY, finP* and *traJ* (Frost *et al.*, 1994). Moreover, *traJ*, the main transfer activator, has the lowest homology among the regulatory genes, sharing only approximately 13% sequence homology with R100 and R1 (Frost *et al.*, 1994). Only the regulatory protein FinO from other F-like plasmids can complement the F *finO* mutation, as FinO produced from a different F-like plasmid can bind *finP* from F and completely restore the fertility inhibition system (Jerome *et al.*, 1999).

B. Regulation of F transfer

The main *tra* operon is activated by the product of the *traJ* gene, TraJ. Other important factors that successfully regulate activation of the *tra* operon are the Histone-like nucleoid structuring protein (H-NS) (Will *et al.*, 2004), the conjugative plasmid expression two-component system (CpxAR) (Gubbins *et al.*, 2002) and the protein SfrA (ArcA) (Silverman *et al.*, 1991).

I. TraJ, the positive regulator

The *traJ* gene is found at position 1033-1719 (from the BglII site in *orf169*) in the transfer region of the F plasmid, just downstream of the *traM* gene and upstream of *traY*, which is the first gene of the long *traYX* operon

(Frost *et al.*, 1994). The TraJ protein is considered to be a 229 amino acid, 27 kDa protein. Through time and further experimentation, TraJ features have been revised and rearranged. This activator was originally thought to be an outer membrane protein because of its insolubility when expressed in large quantities from foreign promoters (Cuozzo *et al.*, 1984). However, its sequence suggested no association with the membrane; moreover, gene activators are rarely found in the membrane. Later, it was shown to be a cytoplasmic protein (Cuozzo & Silverman, 1986).

TraJ activates the *traY* promoter (P_Y) in order to start the production of the transfer machinery; however its actual mechanism of action has proven to be elusive. In the R100 plasmid, TraJ has been shown to bind P_Y *in vitro*, albeit only when fresh and under unusual conditions of low pH (Taki *et al.*, 1998). Its insolubility when overexpressed has also made it difficult to purify in sufficient quantities to test *in vitro*. Surprisingly, its copy number inside the cell is higher (approximately 3000 copies) than most regulators, more so because it is considered to work only on one promoter (Will & Frost, 2006a). Other global regulators, like CRP, have similar numbers in the cell, but they regulate multiple promoters (Raibaud & Schwartz, 1984). TraJ, whose levels remain unchanged in stationary phase, is unable to promote conjugation unless nutritional levels are favorable for further growth (Frost & Manchak, 1998). Being the main activator, TraJ is itself the target of intense regulation: its mRNA is protected from translation when bound to the fertility inhibition antisense RNA *finP* (Sandercock & Frost, 1998); its promoter (P_J) is repressed via a strong silencing mechanism by H-NS from the host (Will *et al.*, 2004); and finally, it is degraded by the host protease-chaperone system HslVU, when the cell is confronted with stress (Lau-Wong *et al.*, 2008).

Since *traJ* is not conserved among IncF plasmids, it has been difficult to distinguish motifs that might shed light on its mechanism of action. Several regulators are able to bind directly to DNA and activate RNA polymerase activity; others activate by interfering with repression. Therefore, one of these two characteristics should be considered to be part of F TraJ's functions, if not both.

II. The Fertility Inhibition (FinOP) system and IS*3*

The fertility inhibition (*finOP*) system consists of two elements, the antisense RNA *finP* and the RNA chaperone protein FinO (van Biesen & Frost, 1994). The *finP* gene encodes a 79-nucleotide antisense RNA which is complementary to the untranslated region of the *traJ* mRNA, including P_J and the ribosome binding site. *finP* regulates the system by binding to the *traJ* mRNA and impeding translation (van Biesen & Frost, 1994). This is called the "kissing-complex" and is achieved by the interaction of two stem loops in the RNA structure, SLI-SLII and SLIc-SLIIc for *finP* and *traJ* mRNA, respectively (Ghetu *et al.*, 2000). *traJ* mRNA forms a third stem loop (SLIII) which is smaller and does not interact with *finP*. There are several different alleles of *finP* that are specific for each F-like *traJ* (Finlay *et al.*, 1986). This variability is the basis for specific inhibition thus, *finP* added *in trans* is incapable of inhibiting *traJ* of a different sub-group (Koraimann *et al.*, 1991). *finP* by itself is unstable and readily degraded by RNase E, having a half-life of approximately 3 minutes (Jerome *et al.*, 1999). FinO, which is a 21.2 kDa, 186 amino acid protein encoded at the distal end of the transfer region of F-like plasmids, is able to bind *finP* and protect it from degradation (Jerome & Frost, 1999). FinO, is also capable of binding to the *traJ* mRNA, which stabilizes the kissing-complex with *finP* (Ghetu *et al.*, 2002). Furthermore, FinO is capable of disrupting the secondary structure of stems in both RNAs, promoting duplexing between

the two (Ghetu *et al.*, 1999). This repression system is highly effective in F-like plasmids; however, it is not present in the F factor because of the insertion of an IS3 in *finO* (Cheah & Skurray, 1986). This lack of FinO, makes *finP* incapable of achieving repression by itself and activation of the *tra* operon is constitutive. F-like plasmids R100-1 and R1-19 also have mutations in *finO* and therefore are also derepressed (Meynell *et al.*, 1968). FinO is highly conserved and only two alleles can be found in the F-like plasmids, type I and type II, classified by the intensity of their repression of mating (Willetts & Maule, 1986). Therefore, if F and a resistance factor (R) were both found in the cell, FinO from the R plasmid would bind to *finP* from F and achieve repression of F transfer. It has also been shown that overexpression of *finP* from a foreign promoter can also lead to repression of F, even in the absence of FinO (Dempsey, 1994).

III. The Histone-like nucleoid structuring protein (H-NS)

The histone-like nucleoid structuring protein is a 15.6 kDa, 136 amino acid global regulator involved mainly in repression of several important and unrelated genes in the cell. Its regulon includes flagella synthesis, acidic stress response, multidrug resistance, virulence, transposition and recombination, horizontal gene transfer, as well as many other genes and operons (Tendeng & Bertin, 2003). H-NS exists in high concentration in the cell (approximately 20,000 copies) and is part of the nucleoid-associated family of regulators (Azam & Ishihama, 1999). These regulators can bind to DNA to alter supercoiling and DNA condensation in a manner similar to histones in eukaryotic systems (Ali Azam *et al.*, 1999). H-NS was shown to bind preferentially to curved DNA and AT-rich sequences in a non-specific way. Since this curvature is commonly found in promoter regions and horizontally transferred genes, H-NS is capable of binding and blocking transcription (Williams & Rimsky, 1997). However, H-NS high-affinity

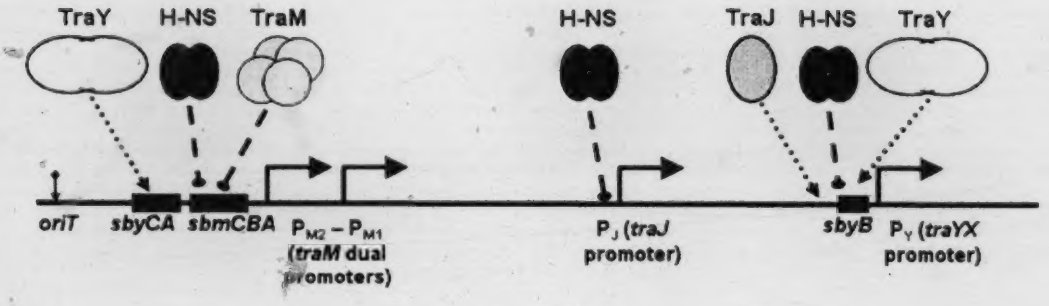
binding sites have been suggested (Bouffartigues *et al.*, 2007), and recent experiments have specified a ten base pair AT-rich consensus sequence for preferred H-NS binding (Lang *et al.*, 2007). The mode of action of H-NS involves protein-protein interactions and oligomerization. H-NS is capable of forming homodimers and higher order oligomers, as well as heterodimers with its paralogue StpA (Suppressor of the *td*⁻ Phenotype A) (Johansson & Uhlin, 1999) as well as with co-repressors such as YmoA (Nieto *et al.*, 2002). This mode of strong repression by H-NS is known as silencing and will be discussed later in the "Gene Silencing" section.

For the F plasmid, H-NS has been shown to react to nutritional stress by inhibiting conjugation as the cell goes into stationary phase (Will *et al.*, 2004). F *traJ* and *traM* both have predicted bends in their promoter regions and genes and gel retardation assays and DNase I footprinting indicate H-NS binding (Will *et al.*, 2004). As illustrated in figure 1.3, *traM, traJ* and *traY* gene expression are directly affected by H-NS in late exponential phase (Will & Frost, 2006a). Furthermore, *hns* mutants show higher levels of TraJ and are capable of maximal conjugative transfer for up to ten hours. This repressive effect occurs only when the *traM* and *traJ* promoters are both present *in cis* (Will *et al.*, 2004). Misleading results showing *traJ* promoter activation in F-like plasmids can be observed using a P_J-*lacZ* fusion, examined in isolation (Starcic-Erjavec *et al.*, 2003). Surprisingly, *hns* mutants do not require TraJ in order to achieve F transfer, which suggests a countersilencing role for TraJ (Will & Frost, 2006a).

IV. Other regulators of F

Many other plasmid- and host-encoded factors regulate F-plasmid expression. Examples of these regulators and their effects are listed below. Other plasmid-encoded regulators include F TraM and TraY, which

Figure 1.3 The effect of H-NS and the three *tra* regulatory proteins on the main promoters of the F transfer region. Dashed lines ending in circles indicate down-regulation, whereas dotted lines ending in arrows indicate activation. Solid rectangles indicate protein-binding sites whereas promoter regions are represented as short bent arrows. The origin of transfer (*oriT*) is represented by a small downward arrow.


TraY
H-NS
TraM
H-NS
TraJ
H-NS
TraY
oriT
sbyCA
sbmCBA
P_M2 – P_M1
(traM dual
promoters)
P_J (traJ
promoter)
sbyB
P_Y (traYX
promoter)

are DNA binding proteins that regulate their own expression (Penfold *et al.*, 1996; Silverman & Sholl, 1996). Moreover, TraY was shown to bind upstream of *traM* and positively regulate its gene expression (Penfold *et al.*, 1996). Environmental regulators include CpxA and SfrA (ArcA); both members of two-component systems. CpxA is a sensor kinase that phosphorylates CpxR when the cell encounters extracytoplasmic stress or alterations to the cell envelope (Raivio & Silhavy, 1997). This stress response activates production of the HslVU protease, which targets and degrades TraJ (Lau-Wong *et al.*, 2008). However, this sensitivity seems to be connected to other factors not yet identified (L. Frost, personal comunication). The ArcA (Aerobic Respiration Control), also known as SfrA (Sex Factor Regulation) was shown to be required for activation of the P_Y promoter in F and R1 plasmids (Strohmaier *et al.*, 1998). ArcA is considered to activate P_Y in concert with TraJ in a mechanism not yet defined (Strohmaier *et al.*, 1998). Other host global regulators include the proteins Hfq (Host Factor for phage Qβ replication) and Dam (DNA adenine methylase). Hfq is a RNA-binding protein that acts as an important post-transcriptional regulator of RNA processes in the cell (Moller *et al.*, 2002). It was recently shown that Hfq represses mating ability by destabilizing both F *traM* and *traJ* mRNA (Will & Frost, 2006b). Hfq binds to a poly-AU region between SLIII and SLIIc of *traJ* mRNA but not to the antisense RNA *finP*, and therefore, does not complement the lack of FinO (Will & Frost, 2006b). Dam methylation is a signaling mechanism that is widely used in *E. coli* (Low *et al.*, 2001). In regard to conjugation, it has been shown to have a positive effect on *finP* production (Torreblanca *et al.*, 1999). Dam methylase has affinity for GATC sites found within P_{finP}, activating the promoter (Torreblanca *et al.*, 1999). However, this modest activation would not be detected in the F plasmid, as the lack of FinO already derepresses the system. Finally, nutritional factors also exert an

effect on conjugation. These factors include Lrp (leucine-responsive regulatory protein) and CRP (cyclic AMP receptor protein). Lrp is a global regulator that acts mainly on genes involved in metabolism in stationary phase (Tani *et al.*, 2002). Its effect on conjugation has been shown for the F-like plasmids pSLT and pRK100 from *Salmonella typhimurium* and uropathogenic *E. coli* (UPEC) respectively (Casadesus & Low, 2006). Lrp binds to the P_J promoter and activates TraJ production, whereas mutants lacking Lrp show a 50- and 4-fold reduction in transfer efficiency for pSLT and pRK100, respectively (Camacho & Casadesus, 2002; Starcic-Erjavec *et al.*, 2003). CRP is a transcriptional factor that directly activates RNA polymerase when in complex with cAMP (Niu *et al.*, 1996). CRP regulates genes involved in catabolism of alternative carbon sources (Zheng *et al.*, 2004). In pRK100 CRP has been shown to bind to the P_J region, upregulating TraJ production (Starcic *et al.*, 2003). Thus, although it remains controversial, starvation was suggested to enhance expression of *traJ* by directly affecting the CRP-cAMP complex in pRK100.

C. Gene silencing

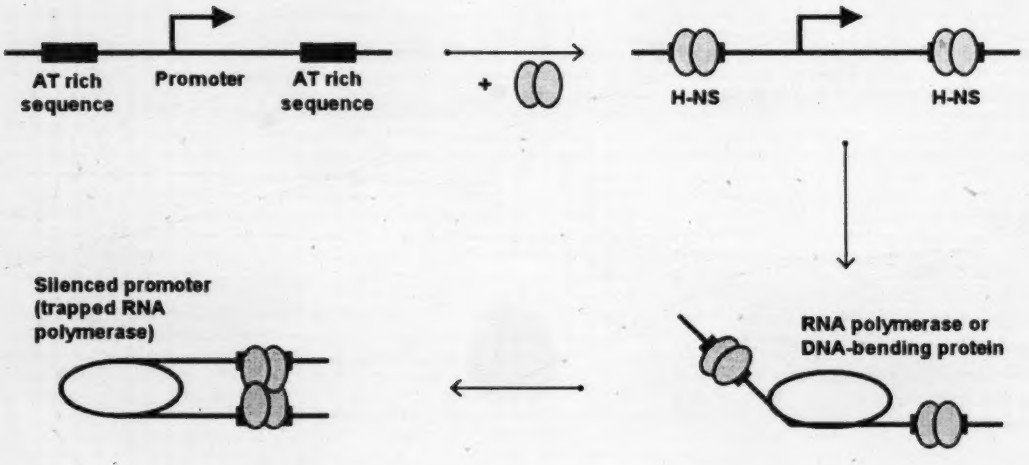
In eukaryotic systems, a large number of genes are not expressed until certain requirements are met (Laurenson & Rine, 1992). Many of these genes are regulated by the topology of the DNA and its superhelical density. In those systems proteins that wrap around and compact DNA, called histones, achieve this repression and are considered to be 'silenced' (Laurenson & Rine, 1992). This kind of DNA-histone complex has been shown to be present in prokaryotes as well. However, regulation of supercoiling and silencing is not achieved by histones, but by histone-like proteins called nucleoid-associated proteins (Ali Azam *et al.*, 1999). H-NS, being a member of this large family of regulators, controls a genome-wide

regulon and also shows an important effect on mobile genetic elements obtained by lateral gene transfer, such as virulence genes and pathogenicity islands (Navarre *et al.*, 2007). However, although they have a similar role to that of histones and are called histone-like, their sequence and structure show almost no homology to eukaryotic histones (Dorman, 2004).

As shown in figure 1.4, H-NS silencing starts when the protein is attracted to a curved region of DNA, generally AT-rich (Navarre *et al.*, 2006). This high-affinity region is called a nucleation site and there are usually two or more nucleation sites flanking a promoter region (Lang *et al.*, 2007). When RNA polymerase binds the DNA to start transcription, the DNA is bent and the nucleation sites come in close proximity. This proximity causes H-NS from the adjacent strands to oligomerize with each other, forming a loop that captures RNA polymerase and impedes gene expression (Shin *et al.*, 2005). H-NS can also nucleate along the DNA and bind to adjacent low-affinity sites in order to completely repress the promoter, in a zipper-like manner. Thus, oligomerization is crucial for proper activity, and silenced regions can be very extensive, including more than one promoter region (Dorman, 2004). Other silencing methods suggest that DNA bending can be initiated by another protein, such as IHF, or even by H-NS itself while nucleating on the DNA. In this model, the zippered promoter would be inaccessible to the polymerase and therefore, be silenced (Nagarajavel *et al.*, 2007).

Silencing has proven to be a fast and effective method of regulating gene expression when environmental conditions are not adequate for growth. Temperature, osmolarity, pH and other environmental factors affect DNA supercoiling, which is the cue for H-NS to regulate expression

Figure 1.4 Silencing of a promoter region by H-NS. H-NS binds to the AT rich sequences flanking the promoter region, then RNA polymerase, or another DNA-bending protein, binds to the promoter region and wraps DNA around itself. Once the two H-NS binding sites are in close proximity with each other, nucleation of H-NS occurs and the promoter is silenced. RNA polymerase can either be trapped in this silenced complex or be unable to access the silenced region.


AT rich sequence
Promoter
AT rich sequence
+
H-NS
H-NS
RNA polymerase or DNA-bending protein
Silenced promoter (trapped RNA polymerase)

of specific genes (Hommais *et al.*, 2001). Moreover, silencing has also been shown to be important in the acquisition of xenogeneic DNA (DNA from foreign sources) (Navarre *et al.*, 2007). Genes obtained from conjugation or phage-mediated transduction can endanger the survival of the cell if they are not regulated. Even genes that would be useful can be a strain for the cell if constitutively expressed. H-NS silencing of xenogeneic DNA is an effective way of avoiding this threat, as most DNA obtained from external elements has a higher AT content than the bacterial genome (Daubin *et al.*, 2003). Genes obtained earlier by lateral gene transfer can usually be detected by their low GC content (Lucchini *et al.*, 2006). Xenogeneic DNA can also be readily degraded by host nucleases, therefore impeding acquisition of useful genes that could improve survival in stressful situations. H-NS binding protects this DNA from degradation and can promote its integration (Navarre *et al.*, 2007). Thus, gene silencing is thought to lead to genome enrichment and bacterial evolution.

I. Mechanisms of H-NS de-silencing

In order to regain expression from silenced promoters, H-NS binding has to be disrupted. There are a number of proposed ways to effectively counteract silencing. The first mechanism involves a factor that binds directly to H-NS disrupting its DNA binding ability, hence removing it from the DNA. An example of these countersilencing factors is a truncated version of H-NS, H-NST, which is found in pathogenic strains of *E. coli* (Williamson & Free, 2005). The truncated version of H-NS lacks the DNA binding domain but is still able to achieve protein-protein interactions. H-NST was shown to bind to H-NS and alleviate repression of the *proU* gene, which is usually found silenced in enteropathogenic *E. coli* (EPEC) (Williamson & Free, 2005).

The second de-silencing mechanism involves a DNA binding protein that competes with H-NS. This activator should be sequence-specific and have high affinity for the binding site. An example of this would be the RovA activator in *Yersinia enterocolitica* and *pseudotuberculosis* (Heroven *et al.*, 2004). RovA is an important activator of the invasin (*inv)* genes and it has also been shown to repress its own expression. Surprisingly, RovA activates *inv* by interacting with the RNA polymerase once it has cleared the promoter region silenced by H-NS. Therefore, RovA is considered to be both a derepressor and an activator of *inv* (Tran *et al.*, 2005).

The third countersilencing mechanism involves an alternative sigma factor that assists RNA polymerase in disrupting silencing. An example of this is RpoS (σ^S or σ^{38}) which is a sigma factor involved in stationary phase gene expression (Shin *et al.*, 2005). It was shown that the *hdeAB* promoter was repressed by H-NS when the polymerase was bound to RpoD, the housekeeping sigma factor (σ^D or σ^{70}). However, expression from this promoter was detected when RpoS was bound to the polymerase instead (Shin *et al.*, 2005). This difference is considered to be due to the different amount of bending that the polymerase promotes in the DNA when bound to different sigma factors. The DNA is not wrapped as tightly when RpoS is bound to the polymerase compared to RpoD. This architectural difference in the DNA arrangement is thought to impede the zipper formation required for H-NS silencing (Shin *et al.*, 2005).

The fourth de-silencing mechanism involves alteration in the promoter DNA topology induced by either environmental factors, like temperature, or by binding of other proteins. An example of this is the *virF* gene, which encodes the main regulator of the invasion operon in *Shigella* (Maurelli & Sansonetti, 1988). The *virF* promoter is repressed solely by H-NS at

temperatures less than 32°C. However when the temperature increases over 32°C, such as body temperature, a change in DNA topology promotes the disruption of H-NS silencing (Tobe *et al.*, 1993). Moreover, this change allows Fis (Factor for inversion stimulation) to bind to the previously silenced region assisting in further activation of *virF* (Prosseda *et al.*, 2004). Therefore, both temperature and an antagonist nucleoid-associated protein work in tandem in order to activate invasion in *Shigella*.

In order to desilence the transfer operon in the F plasmid, one or more of these mechanisms should be considered. Either TraJ has high affinity for DNA and competes for the binding site or it interacts directly with H-NS to disrupt silencing. Alternatively, another nucleoid-associated protein, like IHF or Fis, could be interfering with the H-NS-DNA complex. Moreover, another sigma factor could be binding the polymerase, hence reactivating transcription. This thesis, combined with results from other members of the lab, will help elucidate the mechanisms for *tra* activation.

II. Nucleoid-associated proteins

The members of the family of nucleoid-associated, or histone-like, proteins have an essential role in regulating supercoiling of the DNA in order to maintain the structure of the chromosome in bacteria (Blot *et al.*, 2006). These proteins can promote the formation of segments of DNA with altered coiling or domains (Dorman *et al.*, 1999), which are either inaccessible for transcription, as in promoter silencing, or rendered susceptible to access by the RNA polymerase, as in promoter activation. This dual role of the nucleoid-associated proteins can lead to antagonism between them, as mentioned above for *virF* (Prosseda *et al.*, 2004). The members of this group not only differ in this regard, but they are also different in their expression patterns. These proteins are not expressed at

all times, some are only found in early or late exponential phase whereas some others are found mainly in stationary phase (Ali Azam *et al.*, 1999).

Even though H-NS is the most studied histone-like protein, the family includes several other members such as: IHF, Fis, Lrp, Hfq, StpA, HU (Heat Unstable), Hha/YmoA (Haemolysin/*Yersinia* modulator A), YdgT (YmoA/Hha paralogue), Sfh (*Shigella flexneri* H-NS-like), Dps (DNA-binding Protein from Starved cells), CbpAB (Curved DNA-Binding Protein) and IciA (Inhibitor of Chromosome Initiation) (Azam & Ishihama, 1999). Most of the members of this family are capable of forming heterodimers with each other, either in a cooperative way or as antagonists. As mentioned above, H-NS binds to itself and also to its paralogue StpA to enhance repression (Williams *et al.*, 1996), whereas H-NST binds H-NS to alleviate silencing (Williamson & Free, 2005). Fis and HU have both shown antagonism to H-NS silencing in different systems (Prosseda *et al.*, 2004) (Dame & Goosen, 2002). Moreover, IHF might also have anti-H-NS activity, as it shares homology with HU. H-NS can form heterodimers with both Hha (Nieto *et al.*, 1997) and YdgT (Paytubi *et al.*, 2004). Moreover, Hfq and Lrp seem to bind H-NS or interact with it in a way not completely understood (Muffler *et al.*, 1996; Pul *et al.*, 2005). Some histone-like proteins have the ability to complement for the loss of another, albeit not always completely. As described above, several of these factors have been shown to act in the regulation of the F plasmid; however there could still be more members involved.

D. DNA binding proteins

Most regulators act by directly binding to the DNA in order to either promote RNA polymerase activity or impede it. These activators or

repressors can bind to a specific or nonspecific sequence in the DNA. DNA-binding capabilities are determined by a specific motif that can be located anywhere in the sequence of the protein. Many different motifs have been determined for bacterial regulators. Due to the vast number of different DNA-binding motifs only those related to this study will be explained here.

I. Ribbon-Helix-Helix (RHH)

The RHH motif is comprised of two α-helical domains preceded by a β-sheet, thus the name Ribbon-Helix-Helix (Somers & Phillips, 1992). The β-sheet is essential for the DNA-binding function of the protein and even a single point mutation in the β-sheet can drastically diminish binding (Lum & Schildbach, 1999). Most of the regulators that carry this motif function as homodimers, with their β-strands arranged in antiparallel conformation. This double-ribbon arrangement of the β-strands is responsible for contacting the DNA strand (Somers & Phillips, 1992). The Ribbon-Helix-Helix family of regulators includes proteins such as Arc, Mnt, MetJ, CopG, NikR, AlgZ, ParD and the ω repressor. These proteins are found in a wide array of organisms such as phage, *Salmonella*, *E. coli*, *Pseudomonas* and *Streptococcus sp.* and regulate both chromosomal- and phage/plasmid-encoded genes (Lum *et al.*, 2002).

TraY of the F-plasmid is another example of a RHH DNA-binding protein. This protein has been shown to bind both inside the *oriT* region of the F-plasmid (Site of Binding for TraY: *sbyA*) (Lahue & Matson, 1990) and its own promoter P_Y (*sbyB*) (Nelson *et al.*, 1993), with similar affinity (Figure 1.1). F TraY binds *sbyA* in order to assist in nicking of the plasmid and activation of the P_M promoter, whereas binding to *sbyB* is involved in regulating P_Y expression. However, TraY functions as a monomer instead

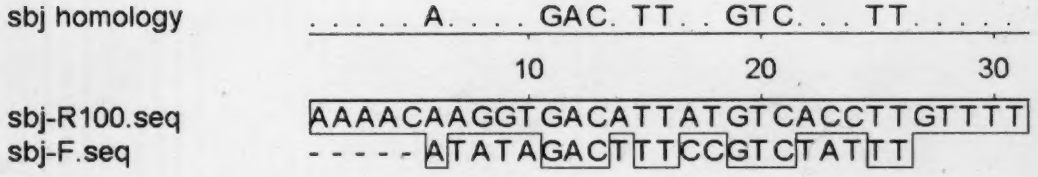
of a dimer (Nelson *et al.*, 1993) because it contains two similar RHH domains in the same protein, thereby functioning as a dimer. This homology between its N-terminus and C-terminus suggests that F TraY was a result of gene duplication (Anthony *et al.*, 1999). Other F-like plasmids, as in the case of R100, have a smaller version of TraY of approximately half the size. This F-like TraY has only one RHH domain; therefore it must form dimers in order to achieve proper DNA-binding (Inamoto *et al.*, 1988).

II. Helix-Turn-Helix (HTH)

The HTH motif is comprised of two α-helices containing a short curved segment between them, thus the name Helix-Turn-Helix (Steitz *et al.*, 1982). The second α-helix is considered to interact directly with the major groove of the DNA strand, binding to it. This motif generally shows low sequence specificity, however structural homology of this domain is high throughout the different proteins carrying this motif (Steitz *et al.*, 1982). This domain can be anywhere from 18 to 25 amino acids long and usually includes a glycine at the turn (Pabo & Sauer, 1984). The members of this family of regulators also function as dimers, with some of them being able of forming higher order oligomers. Many bacterial proteins carry this motif, including: cro repressor, CAP, IS*3* isomerases, *lac* and *gal* repressors, LuxR, FixJ, λcI-II, LysR, just to name a few examples (Ohlendorf *et al.*, 1983).

Although, TraJ binding to DNA has not been shown for F TraJ, it has been reported for F-like R100 TraJ (Taki *et al.*, 1998). The R100 TraJ binding site, *sbj*, was shown to be in an inverted repeat upstream of *traY*, a site that is also found in F (Figure 1.5) and throughout F-like P_Y sequences (Taki *et al.*, 1998). Regulators carrying a HTH motif generally bind to

Figure 1.5 Sequence comparison between the inverted repeat of the *traJ-traY* intergenic region of R100 and F.


sbj homology
. . . . . A. . . . GAC. TT. . GTC. . . TT. . . . .
10
20
30
sbj-R100.seq
AAAACAAGGTGACATTATGTCACCTTGTTTT
sbj-F.seq
- - - - - ATATAGACTTTCCGTCTATTT

inverted repeat sequences (dyad symmetry). These results support the presence of a HTH DNA-binding motif in R100 TraJ. F TraJ binding to the DNA could then activate P_Y by either interacting with RNA polymerase, or by competing for binding sites with H-NS in order to prevent silencing. However, actual existence of this domain has not been proven up to date for any TraJ. In addition, the lack of a crystal structure of TraJ has made its overall architecture difficult to determine.

III. Winged-helix domain

The winged-helix motif was originally observed in hepatocyte nuclear factor three (Lai *et al.*, 1990) and in the fork head mutant proteins of *Drosophila* (Lai *et al.*, 1993). Shortly after, it was observed in regulatory proteins in prokaryotic systems as well (Martinez-Hackert & Stock, 1997). The winged-helix motif is a complex DNA-binding domain that is related to the HTH motif. However, the 'turn' in this motif is significantly longer than the one in the HTH motif. Two loops and two β-strands usually located at the C-terminus of the protein flank the main α-helix, which is closer to the N-terminus. This arrangement around the α-helix, reminiscent of butterfly wings, gives its name to this motif (Gajiwala & Burley, 2000). However, variants of this motif have also been reported for other regulators, therefore this domain is flexible to some extent (Ellison & Miller, 2006). Proteins carrying this motif have been shown to form homodimers and heterodimers, as well as occurring as monomers (Martinez-Hackert & Stock, 1997).

The regulator of virulence (RovA), shown to autoregulate itself and countersilence H-NS, includes a winged-helix motif (Tran *et al.*, 2005). This domain in RovA is a variant that involves a HTH motif followed by two β-strands. RovA is part of the MarR/SlyA family of regulators, involved

mainly in regulation of virulence genes (Heroven *et al.*, 2004). Regulators from this group are found in *Pseudomonas, Serratia, Salmonella, Yersinia, Erwinia, Vibrio and Staphylococcus* species, as well as *E. coli*. As described above, these proteins share low sequence homology between them, albeit maintaining structural homology in their winged-helix domain (Ellison & Miller, 2006).

E. Research objectives

The main objective of this study is to demonstrate the mechanism of action of F TraJ, including whether it can bind to DNA, namely the *traY* promoter region. TraJ binding to DNA could then impede H-NS silencing or counteract it. This objective will be pursued by immuno-precipitation *in vivo*, as insolubility of the TraJ protein makes its difficult to purify it in enough quantities to test *in vitro*. Moreover, the putative HTH of F TraJ will be tested by performing site-directed mutagenesis of conserved amino acids throughout F TraJ. Complementation assays of an amber mutant of TraJ in F (F*lactraJ90*) will be performed in order to determine the activation of mating ability in these mutants.

A secondary objective will involve determination of the oligomeric state of F TraJ by cross-linking of wild-type F TraJ, as most proteins carrying a HTH site act as dimers. Crosslinked protein will be separated by SDS-PAGE and observed by western transfer. Additional TraJ mutants will be generated in order to test the importance of the C-terminal tail of TraJ. In addition, mass spectrometry (MS) will be performed on polyacrylamide gel fragments containing F TraJ in order to detect any modifications that would alter its size in stationary phase. The MS data also revealed that the

reported start codon was incorrect, which was the basis for testing whether the methionine at position 4 was the true start codon.

Chapter 2. Materials and Methods

A. *E coli* strains and plasmids used

Strains used in this study are listed in Table 2.1, whereas plasmids used are listed in Table 2.2. Oligonucleotides used in this work are listed in Table 2.3. Most point mutants were made in pBAD24TraJ (Gubbins *et al.*, 2002) or ligated into the pBAD24 vector (NcoI-HindIII). pCR4Blunt-TOPO (Invitrogen), was used to produce the C-terminal mutants by blunt ligation of the *traJ* PCR products. These PCR fragments were removed from pCR4Blunt-TOPO and inserted into pBAD24 (NcoI-HindIII). Plasmid pILJ11 is a construct derived from the commercial vector pBCSK+ (Stratagene) and includes the F *traJ* gene with its respective upstream and downstream intergenic regions, as well as C- and N-terminal fragments of *traM* and *traY* respectively (SalI-PstI 1586 bp fragment). pILJ11 was used in order to produce the point mutants of the first and fourth methionine in *traJ*, in order to express *traJ* from its own promoter rather than a foreign one such as in pBAD24. The plasmid pILJ14Δ6 transformed in PD32 was cloned from the ClaI-HindIII fragment of the mutant pJMRΔ6 in the vector pBAD33. Oligonucleotides used for mutagenesis were 30 or 39 base pairs (bp) long depending on whether one or more base pairs were mutated. All mutants were sequenced in the University of Alberta Molecular biology facility (MBSU) using a DYEnamic ET sequencing kit (Amersham Biosciences) on a 3730 DNA analyzer (Applied Biosystems).

Table 2.1 List of *Escherichia coli* K-12 strains used.

Strain	Relevant Genotype	Reference
DH5α	F⁻ *hsdR17 ΔlacU169* (Φ80 *lacZΔM15*)*recA1 supE44 endA1 gyrA96 relA1 thi-1*	(Hanahan, 1983)
ED24	F⁻ Lac⁻ Sp^r	(Achtman *et al.*, 1971)
MC4100	F⁻ *araD139* Δ(*argF-lac*)*U169 rpsL150*(Str^r) *relA1 flb5301 deoC1 ptsF25 rbsR*	(Silhavy *et al.*, 1984)
PD32	MC4100 *hns206*::Ap^r	(Dersch *et al.*, 1993)

Table 2.2 List of plasmids used.

Plasmid	Genotype	Parental Plasmid	Reference
F*lac*	F'*lac* Tra$^+$ *finO*$^-$	F	(Achtman *et al.*, 1971)
F*lac traJ90*	F'*lac* Tra$^-$ *finO*$^-$ *traJQ26amber*	F	(Achtman *et al.*, 1971)
pBAD24	Ap^r P_{BAD}	-	(Guzman *et al.*, 1995)
pBAD24TraJ	P_{BAD} *traJ*	pBAD24	(Gubbins *et al.*, 2002)
pJMR34	*traJG169R*	pBAD24	This work
pJMR512	*traJG169A*	pBAD24	This work
pJMR56	*traJY166D*	pBAD24	This work
pJMR534	*traJY166A*	pBAD24	This work
pJMR78	*traJH172D*	pBAD24	This work
pJMR556	*traJH172A*	pBAD24	This work
pJMR245	*traJD5A*	pBAD24	This work
pJMR423	*traJQ14K*	pBAD24	This work
pJMR267	*traJP31A*	pBAD24	This work
pJMR445	*traJS65A*	pBAD24	This work
pJMR467	*traJE77A*	pBAD24	This work
pJMR212	*traJW118A*	pBAD24	This work
pJMR367	*traJI181A*	pBAD24	This work
pJMR389	*traJS186A*	pBAD24	This work
pJMR690	*traJC224A*	pBAD24	This work
pJMR489	*traJI225L*	pBAD24	This work
pJMR712	*traJN227A*	pBAD24	This work
pJMR678	*traJC224ochre*	pBAD24	This work
pJMR756	*traJI225ochre*	pBAD24	This work
pJMR812	*traJI226ochre*	pBAD24	This work
pJMR645	*traJR229ochre*	pBAD24	This work
pJMRΔ30	*traJΔ(200-229)*	pBAD24	This work
pJMRΔ15	*traJΔ(214-229)*	pBAD24	This work

Table 2.2 continued...

Plasmid	Genotype	Parental Plasmid	Reference
pJMRΔ10	*traJΔ(220-229)*	pBAD24	This work
pJMRΔ6	*traJΔ(224-229)*	pBAD24	This work
pJMRΔ1	*traJΔ229*	pBAD24	This work
pJMRΔ3N	*traJΔ(1-3)*	pBAD24	This work
pCR4Blunt-TOPO	Ap^r Km^r P_{lac} *lacZ::ccdB*	-	Invitrogen
pBAD33	Cm^r P_{BAD}	-	(Guzman *et al.*, 1995)
pILJ14	P_{BAD} *traJ*	pBAD33	(Lau-Wong *et al.*, 2008)
pILJ14Δ6	*traJΔ(224-229)*	pBAD33	This work
pBCSK+	Cm^r	-	Stratagene
pILJ11	*traJ*	pBCSK+	(Lau-Wong, unpublished)
pJM1T	*traJM1T*	pBCSK+	This work
pJM4T	*traJM4T*	pBCSK+	This work
pJM1-4T	*traJM1T, traJM4T*	pBCSK+	This work

Table 2.3 List of oligonucleotides used.

Oligonucleotide	Sequence	Use
A2426	5'AAACACGCATCTCTGATATGCGAC3'	Sequencing of *traJ*
LFR181	5'CAGGCTGAAAATCTTCTC3'	pBAD24 Reverse
LFR182	5'TACTGTTTCTCCATACCC3'	pBAD24 Forward
MRO3	5'TACTTATATAGCACC-CGAATCAGTCATAAT3'	pJMR34
MRO4	5'ATTATGACTGATTCG-GGTGCTATATAAGTA3'	(G169R)
MRO5	5'GATGTCTTTTACTTA-GATAGCACCGGAATC3'	pJMR56
MRO6	5'GATTCCGGTGCTATC-TAAGTAAAAGACATC3'	(Y166D)
MRO7	5'AGCACCGGAATCAGT-GATAATGCTATAGCA3'	pJMR78
MRO8	5'TGCTATAGCATTATC-ACTGATTCCGGTGCT3'	(H172D)
MRO10	5'TGGCAATTTTATGATGCTGCTCAT3'	Sequencing of *traJ*
MRO17	5'AAGCTTTTAAACAGAGAAGT-AGTCGCATATCAGAGATGC3'	pJMRΔ30 (*traJ*Δ30)
MRO21	5'AGAAATATTATTCTG-GCGCAATTTTATGAT3'	pJMR212
MRO22	5'ATCATAAAATTGCGC-CAGAATAATATTTCT3'	(W118A)

Table 2.3 continued...

Oligonucleotide	Sequence	Use
MRO23	5'AAGCTTTTAAAACTTTTTAT TGTAGAGTAAGATAATTAA3'	pJMRΔ15 (*traJ*Δ15)
MRO24	5'ATGTATCCGATGGCT CGTATTCAACAAAAA3'	pJMR245
MRO25	5'TTTTTGTTGATTACG AGCCATCGGATACAT3'	(D5A)
MRO26	5'CAGGATTATCCAAAT GCAGCCTGTATCAGG3'	pJMR267
MRO27	5'CCTGATACAGGCTGC ATTTGGATAATCCTG3'	(P31A)
MRO36	5'GCAAGAATATTAAAT GCATCCATCTCCACA3'	pJMR367
MRO37	5'TGTGGAGATGGATGC ATTTAATATTCTTGC3'	(I181A)
MRO38	5'ATATCCATCTCCACA-GCAAAGAAACACGCA3'	pJMR389
MRO39	5'TGCGTGTTTCTTTGC-TGTGGAGATGGATAT3'	(S186A)
MRO42	5'CAAAAACATGCTCGT AAAATAGATCGTCTG3'	pJMR423
MRO43	5'CAGCAGATCTATTTT ACGAGCATGTTTTTG3'	(Q14K)
MRO44	5'GAAAAATGGCTTCTG GCGCAGAGAGATTTT3'	pJMR445
MRO45	5'AAAATCTCTCTGCGC CAGAAGCCATTTTTC3'	(S65A)

Table 2.3 continued...

Oligonucleotide	Sequence	Use
MRO46	5'TTGATCTCTGTCACA GCGATGGAAGCATAT3'	pJMR467
MRO47	5'ATATGCTTCCATCGC TGTGACAGAGATCAA3'	(E77A)
MRO48	5'GAGAAGGCTATGTGT CTCATAAATACGCGT3'	pJMR489
MRO49	5'ACGCGTATTTATGAG ACACATAGCCTTCTC3'	(I225L)
MRO50	5'AAGCTTTTAGTATAAATAAT AAATAAACTTTTTATTGTA3'	pJMRΔ10 (*traJ*Δ10)
MRO51	5'TACTTATATAGCACC-GCAATCAGTCATAAT3'	pJMR512
MRO52	5'ATTATGACTGATTGC-GGTGCTATATAAGTA3'	(G169A)
MRO53	5'GATGTCTTTTACTTA-GCTAGCACCGGAATC3'	pJMR534
MRO54	5'GATTCCGGTGCTAGC-TAAGTAAAAGACATC3'	(Y166A)
MRO55	5'AGCACCGGAATCAGT-GCTAATGCTATAGCA3'	pJMR556
MRO56	5'TGCTATAGCATTAGC-ACTGATTCCGGTGCT3'	(H172A)
MRO57	5'AAGCTTTTACATAGCCTTCT-CGTATAAATAATAAATAAA3'	pJMRΔ6 (*traJ*Δ6)
MRO59	5'CCATGGATCGTATTCAACA-AAAACATGCTCGTCAAATA3'	pJMRΔ3N (*traJ*Δ3-N)

Table 2.3 continued...

Oligonucleotide	Sequence	Use
MRO60	5'TGTGAGGAGGTTCCT-ACGTATCCGATGGAT3'	pJM1T
MRO61	5'ATCCATCGGATACGT-AGGAACCTCCTCACA3'	(M1T)
MRO62	5'TATCCGACGGATCGT-ATTCAACAAAAACAT3'	pJM4T & pJM1-4T
MRO63	5'ATGTTTTTGTTGAAT-ACGATCCGTCGGATA3'	(M4T & M1T/M4T)
MRO64	5'TACGAGAAGGCTATGTGTATCAT AAATACGTAATAAGGTGTTAAT3'	pJMR645
MRO65	5'ATTAACACCTTATTACGTATTTA TGATACACATAGCCTTCTCGTA3'	(R229ochre)
MRO67	5'TTATACGAGAAGGCTATGTA-AATCATAAATACGCGTTAA3'	pJMR678
MRO68	5'TTAACGCGTATTTATGATTT-ACATAGCCTTCTCGTATAA3'	(C224ochre)
MRO69	5'TTATACGAGAAGGCTATGGC-TATCATAAATACGCGTTAA3'	pJMR690
MRO70	5'TTAACGCGTATTTATGATAG-CCATAGCCTTCTCGTATAA3'	(C224A)
MRO71	5'AAGGCTATGTGTATCATAGC TACGCGTTAATAAGGTGTT3'	pJMR712
MRO72	5'AACACCTTATTAACGCGTAG CTATGATACACATAGCCTT3'	(N227A)
MRO75	5'GAGAAGGCTATGTGTTAAAT-AAATACGCGTTAATAAGGT3'	pJMR756
MRO76	5'ACCTTATTAACGCGTATTTA-TTTAACACATAGCCTTCTC3'	(I225ochre)

Table 2.3 continued...

Oligonucleotide	Sequence	Use
MRO81 MRO82	5'GAGAAGGCTATGTGTATCTA AAATACGCGTTAATAAGGT3' 5'ACCTTATTAACGCGTATTTT AGATACACATAGCCTTCTC3'	 pJMR812 (I226ochre)
RWI91	5'GCGACTACTTCTCTGTTTCTAATAA3'	ChIP PCR (*1613-1637)
RWI92	5'CCCGCTGTTTATCTTCTGGTTACC3'	ChIP PCR (*1812-1788)
traC-5	5'GTGAATAACCCACTTGAGGCCG3'	ChIP PCR
traC-3	5'TCATGCCACACTCCTGTATTTCTC3'	ChIP PCR

*Position in (Frost *et al.*, 1994)

B. Chemicals, growth media and enzymes

Cultures were grown in Luria Bertani broth (LB: 1% Difco triptone, 0.5% Difco yeast extract, 1% NaCl) and arabinose (Sigma) was used to induce *traJ* in pBAD constructs. Colonies were grown on LB agar (1.5% Difco Bacto agar) or 2% water agar (5 mM $MgSO_4$, 1.5% Difco Bacto agar, double distilled (dd) H_2O) with 1xM9 minimal salts (5 mM $Na_2HPO_4 \cdot 7H_2O$, 22mM $KHPO_4$, 8 mM NaCl, 19 mM NH_4Cl) and 0.4% lactose. Antibiotic final concentrations were as follows: ampicillin (Ap) 100 µg/ml, chloramphenicol (Cm) 20 µg/ml, kanamycin (Km) 25 µg/ml, streptomycin (Sm) 200 µg/ml and spectinomycin (Sp) 100 µg/ml. Taq polymerase, restriction enzymes and their buffers were purchased from Fermentas. PfuTurboAD polymerase and buffer were purchased from Stratagene. Formaldehyde was purchased from Anachemia, whereas disuccinimidyl suberate (DSS) was purchased from Pierce. Protein A beads were purchased from Roche. DNA (QIAprep Spin Miniprep Kit) and gel extraction (QIAquick GEL Extraction Kit) solutions and columns were purchased from Qiagen.

C. Site-directed mutagenesis

The protocol used for site-directed mutagenesis was based on the one described by Stratagene (Quickchange site-directed mutagenesis Kit handbook) with small changes. In order to produce pJMR34, 50 ng of pBAD24TraJ was added to a mixture of 10x PfuTurboAD reaction buffer (Stratagene), 0.2 mM dNTP mix, 125 ng of each forward (MRO3) and reverse (MRO4) mutagenic primers and ddH_2O to a final volume of 50 µl. Oligonucleotides used here were previously purified by electrophoresis on an 8% polyacrylamide gel containing urea (8 M). PfuTurbo AD polymerase

(1 µl) was then added to the mixture and PCR was performed (95°C for 30 sec, 55°C for 1 min, 68°C for 10 min, 18 total cycles). Samples were treated with Dpn1 for two hours at 37°C in order to digest the methylated parental DNA. Finally, the 5.2 kb band was purified from an agarose gel (1%) and transformed into rubidium chloride competent DH5α cells. Cells were plated on LB agar with ampicillin and incubated overnight at 37°C. Transformants were grown in LB broth, pJMR34 was extracted, sequenced (MRO10) and transformed into MC4100/F*lac traJ90* to test for complementation of mating ability. All other point-mutations were made in the same manner with their respective mutagenic oligonucleotides (listed in Table 2.3). Template DNA used to produce the methionine mutants was pILJ11 instead of pBAD24TraJ.

D. Deletion mutagenesis

In order to produce pJMRΔ30, oligonucleotide MRO17 was designed and used in combination with LFR182, an oligonucleotide for the pBAD24 multiple cloning site immediately upstream of *traJ* (including the NcoI site used to clone *traJ* into pBAD24). MRO17 targeted the C-terminal region of *traJ* inserting a HindIII restriction site immediately after a stop codon (TAA) in position 200 of the amino acid sequence of *traJ*. After PCR, the 600 bp *traJ*Δ30-HindIII fragment was purified on a 1% agarose gel. Then, 20 ng of the PCR fragment was mixed with 0.5 µl salt solution (1.2 N NaCl, 60 mM $MgCl_2$), 2.5 µl ddH2O and 0.5 µl pCR4BluntTOPO as indicated by Invitrogen. The reaction was incubated at room temperature (25°C) for five minutes and transformed into *E. coli* One-shot competent cells (Invitrogen). The mixture was plated on LB agar with ampicillin or kanamycin. Plasmid DNA was extracted from transformed colonies and was digested at 37°C with NcoI and HindIII; pBAD24 was also digested in the same manner. The

DNA fragments were separated on a 1% agarose gel and the 600 bp NcoI-HindIII *traJΔ30* band was excised, as well as the pBAD24 vector. Both gel bands were co-purified and ligated overnight with T4 DNA ligase at 16°C. The ligation was transformed into rubidium chloride competent DH5α cells and plated on LB agar with ampicillin. pJMRΔ30 was then extracted, sequenced (A2426) and transformed into MC4100/F*lac traJ90* cells to test for complementation. All other deletion mutants were produced in the same manner with their respective mutagenic oligonucleotide (listed Table 2.3). Because the *traJΔ3N* mutation was in the N-terminus, LFR181 was used instead of LFR182 to produce pJMRΔ3N.

E. Mating assays

MC4100/F*lac traJ90* containing pBADTraJ or mutants derived from it were used as donor strains and ED24 was used as the recipient. Overnight cultures of both donor and recipient cells (0.15 ml), were used to inoculate 3 ml of LB broth with no antibiotics. Cultures were grown at 37°C and after one hour 0.1% arabinose was added to donors in order to induce TraJ production. At OD_{600} of approximately 1.0, 100 µl of both donor and recipient cells were incubated together for one hour at 37°C in 1 ml of fresh LB with 0.1% arabinose. Samples were vortexed for 5 seconds to disrupt mating and serial dilutions were made with 1x SSC (150 mM NaCl, 15 mM sodium citrate, pH 7.0) for each test reaction. From each dilution (10^{-2} to 10^{-6}), 10 µl was placed on agar plates with streptomycin and ampicillin to select for donors, and minimal media with 1xM9 minimal salts, lactose and spectinomycin to select for transconjugants. Finally, colonies were counted and mating efficiency was determined as a ratio of transconjugants/100 donors. Simultaneously, another 100 µl aliquot was sampled and cells were pelleted for immunoblot analysis.

F. Immunoblot analysis

Cell pellets were resuspended in 10 µl loading buffer (0.5 M Tris-HCl pH 6.8, 1 M dithiothreitol (DTT), 10% sodium dodecyl sulphate (SDS), 1% Bromophenol blue, 10% glycerol) (Laemmli, 1970) and heated in a tempblock (95°C) for five minutes. Samples were electrophoresed on a 12% polyacrylamide resolving gel (30% acrylamide, 0.8% bis-acrylamide, 0.05% tetramethyl-ethyenediamine (TEMED), 0.1% ammonium persulfate) at 200 mV for 50 minutes in protein gel running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS). Protein was then transferred (100 mV for one hour) onto an Immobilon-P membrane (Millipore) in Towbin buffer (25 mM Tris, 192 mM glycine, 20% methanol, 0.004% SDS) (Towbin *et al.*, 1979). The membrane was incubated in 5% skim milk powder in TBST (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween 20) overnight. Membranes were further incubated one hour in the same solution with polyclonal anti-TraJ antiserum (1:20,000 dilution). After incubation, membranes were washed four times in TBST at room temperature for ten minutes each. Membranes were then incubated for one more hour with 5% skim milk powder solution containing 1:20,000 dilution of secondary antibody (horseradish peroxidase-conjugated donkey anti-rabbit immunoglobulin G, GE Health Care), and washed four times as before. Finally, the membrane was developed with Western Lightning Chemiluminescence Reagent Plus (PerkinElmer Life Sciences) and exposed to BioMax Xar film (Kodak).

G. Chromatin Immuno-precipitation (ChIP) assay

The protocol for the ChIP assay was based on the methods used by Strahl-Bolsinger (Strahl-Bolsinger *et al.*, 1997) with some changes. First 50

ml of LB broth were inoculated with 2.5 ml of overnight cultures and incubated at 37°C with shaking. At OD_{600} of approximately 1.0, 1% formaldehyde was added and samples were incubated 15 minutes at room temperature. Glycine (125 mM) was added and the mixture was incubated five minutes at room temperature. Cells were washed once with phosphate-buffered saline (PBS) (1.5 mM KH_2PO_4, 137 mM NaCl, 8 mM Na_2HPO_4, ddH_2O, pH 7.2) and resuspended in 300 µl ChIP lysis buffer (50 mM HEPES pH 7.5, 140 mM NaCl, 1% Triton X100, 0.1% Sodium deoxycholate, plus one Complete-Mini protease inhibitor cocktail tablet (Roche)). Glass beads were then added and the mixture was incubated with shaking for 30 minutes at 4°C at maximum level. Samples were centrifuged and the liquid fraction was resuspended and sonicated for 30 seconds (3.5 peak-peak amplitude, Branson Sonifier 450). Samples were pelleted (4°C) and supernatants were transferred to new tubes. A 15 µl aliquot of each sample was set aside at this point to use later as Total DNA samples. Anti-TraJ antibodies (1:20,000) were added to the supernatants and the mixtures were incubated for two hours with shaking at 4°C. Protein A beads (50 µl), previously washed with ChIP lysis buffer, were added and incubated for four hours with shaking at 4°C. Immuno-precipitations were washed twice with 1 ml ChIP lysis buffer, twice with 1 ml ChIP high-salt lysis buffer (ChIP lysis buffer with 500 mM NaCl), twice with 1 ml ChIP wash buffer (10 mM Tris pH 8, 250 mM LiCl, 0.5% NP-40, 0.5% NaDeoxycholate, 1 mM EDTA) and twice with 1 ml TE buffer. Samples were eluted by adding 75 µl elution buffer (50 mM Tris pH 8, 1% SDS, 10 mM EDTA) twice and the beads were incubated ten minutes at 65°C. Both elutions were combined and incubated overnight at 65°C. Samples were then purified using the Qiagen PCR purification kit and eluted with 50 µl ddH_2O. Different concentrations of DNA were used for each 25 µl PCR reaction, using primers specific for the promoter region of *traY* (RWI91-

RWI92), to produce a 200 bp fragment. PCR products were run on a 1.5% agarose gel and stained with ethidium bromide. Finally, the bands were observed in UV light and photographed.

H. Mass spectrometry

In preparation for this experiment, all glassware was thoroughly cleaned with 0.1 N HCl. Moreover, every solution used was filtered through a 0.45 µm membrane (Nalgene), including the 30:0.8 acrylamide solution, the electrophoresis running buffer, Commassie blue stain (50% methanol, 10% acetic acid, 0.05% Serva Blue R-250, 0.05% Serva Blue G) and destain solutions (10% acetic acid, 20% methanol).

A tube with 3 ml of LB broth (with Sm and Ap) was inoculated with 0.15 ml of an overnight culture of MC4100/F*lac traJ90*/pBAD24TraJ. MC4100/F*lac traJ90*/pBAD24 was inoculated as a negative control. Cultures were incubated at 37°C for one hour and then induced with arabinose (0.1%). Cultures were further incubated for two and four hours. Pellets at 0.2 OD_{600} were obtained for each time point and boiled at 95°C for 10 minutes in 10 µl loading buffer. Samples were run on a 15% acrylamide gel (200 mV, 50 minutes) and the gel was stained with Commassie blue for 10 minutes and destained for 20 minutes. The induced protein band, representing TraJ, was cut from the gel using a sterile scalpel and stored in 100 µl ddH_2O. The gel fragment was sent to the Institute for Biomolecular Design (IBD) in the Biochemistry department at the University of Alberta in order to perform mass spectrometry with an Applied Biosytems Voyager system using the ESI-TOF (electrospray ionization time-of-flight) technique.

I. Cross-linking with DSS

The protocol for the cross-linking was based on the methods used by Klimke (Klimke *et al.*, 2005) with some changes. Three ml cultures (LB broth with 0.1% arabinose) were inoculated with 0.15 ml of MC4100/F*lac traJ90*/pBAD24TraJ overnight cultures until the OD_{600} was approximately 1.0. Pellets from two 1 ml aliquots from were obtained per culture, which were washed three times with PBS buffer and resuspended in 200 µl PBS. One sample was cross-linked with 0.5 mM DSS, whereas the second was used as a negative control. Reactions were incubated for 30 minutes at room temperature and then quenched by adding 60 mM Tris pH 7.5 and incubating five more minutes. Cells were pelleted and washed with cold PBS and resuspended in 50 µl Towbin buffer. Samples were then boiled for five minutes and electrophoresed on a 12% polyacrylamide SDS gel. Finally, TraJ was detected by immunoblot, as described above. The H-NS mutant strain (PD32) used to test for oligomerization is Ap^r, thus pILJ14 (Cm^r) was used instead of pBAD24TraJ.

Chapter 3. TraJ is a DNA binding protein with an essential four amino acid C-terminal tail

A. Introduction

TraJ is considered to be the main activator of the transfer operon, which encodes most of the genes required for F conjugation. Although its mechanism for activation is not completely determined, it has been shown to bind the *traYX* promoter (P_Y) in the F-like R100 plasmid (Taki *et al.*, 1998). Although the two systems are similar, DNA binding has not been shown for F TraJ. Moreover, the dyad sequence required for TraJ binding in R100, *sbj*, is also found in F, although the sequence is not conserved, suggesting specificity of binding (Frost *et al.*, 1994). Due to the low solubility of F TraJ, it has been difficult to purify enough of it to test its DNA-binding ability *in vitro*. Moreover, TraJ is unstable and loses activity almost immediately after purification. DNA-binding tests using R100 TraJ required fresh protein (less than 24 hours) to avoid inactivation of the protein, even when the protein was frozen (Taki *et al.*, 1998).

In this study, gene alignment of TraJ from different F-like plasmids, including R100, was performed in order to identify a putative motif for DNA-binding in F TraJ. Mutagenesis of a selection of conserved amino acids was then performed in *traJ* cloned into pBAD24 and the resultant mutants were transformed into MC4100/F*lac traJ90* (F^- phenotype) and tested for mating ability. The presence of the protein was determined by immunoblot analysis with anti-TraJ antibodies. F TraJ deletion mutants at the C-terminus were also produced in order to determine whether the C-terminal region was essential for activity. Finally, binding of F TraJ to P_Y was tested *in vivo* using the Chromatin Immuno-precipitation (ChIP) assay, which involves cross-linking of a DNA-protein complex, immuno-precipitation and PCR on the immuno-precipitated DNA fragments to

detect the P_Y region. Samples were also taken at different time points to determine whether TraJ was able to bind DNA in stationary phase.

B. Results

I. Sequence alignment of TraJ from F-like plasmids and determination of a putative DNA-binding site.

The sequence of F TraJ was analyzed using the BLAST program (NCBI) in order to find its closest relatives. Two TraJ proteins from different conjugative F-like plasmids had the highest homology scores: R100, R1-19 and pOU127. Seven TraJ-like proteins from different strains and plasmids were also identified in this manner. Sequence alignment was performed (MegAlign, DNAstar), as shown in Figure 3.1 and 3.2. The sequence identity was strikingly low, and very few conserved amino acids were found amongst TraJ from the different sources. However, two clusters of conserved amino acids were identified, one closer to the N-terminus (residues 30-55) and another closer to the C-terminus (residues 160-185). From previous reports, it has been suggested that the TraJ protein includes a Helix-Turn-Helix (HTH) site in its sequence (Frost *et al.*, 1994). Furthermore, a HTH motif is usually around 20-25 amino acids in length as are these conserved domains, and it usually includes a conserved glycine, as does the second cluster. Therefore, the second cluster was considered to be the putative HTH DNA-binding site and mutations were planned for the conserved amino acids found in this region.

Figure 3.1 Sequence alignment of F TraJ (Accession: BAA97942) and two orthologues from plasmids R100 (Accession: BAA78849) and R1-19 (Accession: AAA92657). Residues that match F TraJ sequence are shown on top of the alignment as tall bars. A ruler is set in reference to F TraJ and grey areas refer to alignment adjustments in the F TraJ sequence. Boxed residues are conserved amino acids compared to the F TraJ sequence. Two clusters of conserved amino acids can be observed, one between residues 29 to 52 and between residues 162 to 186, as indicated by double-headed arrows.

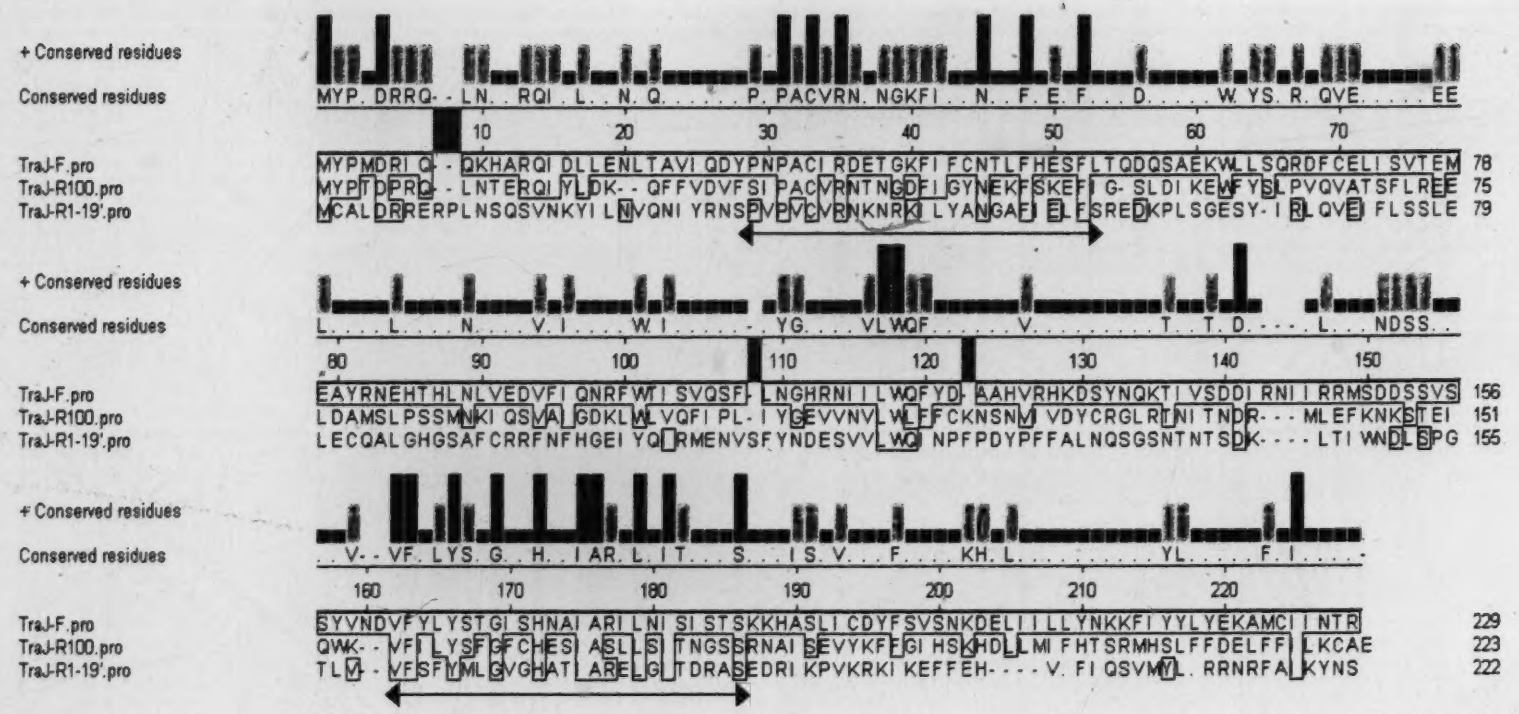
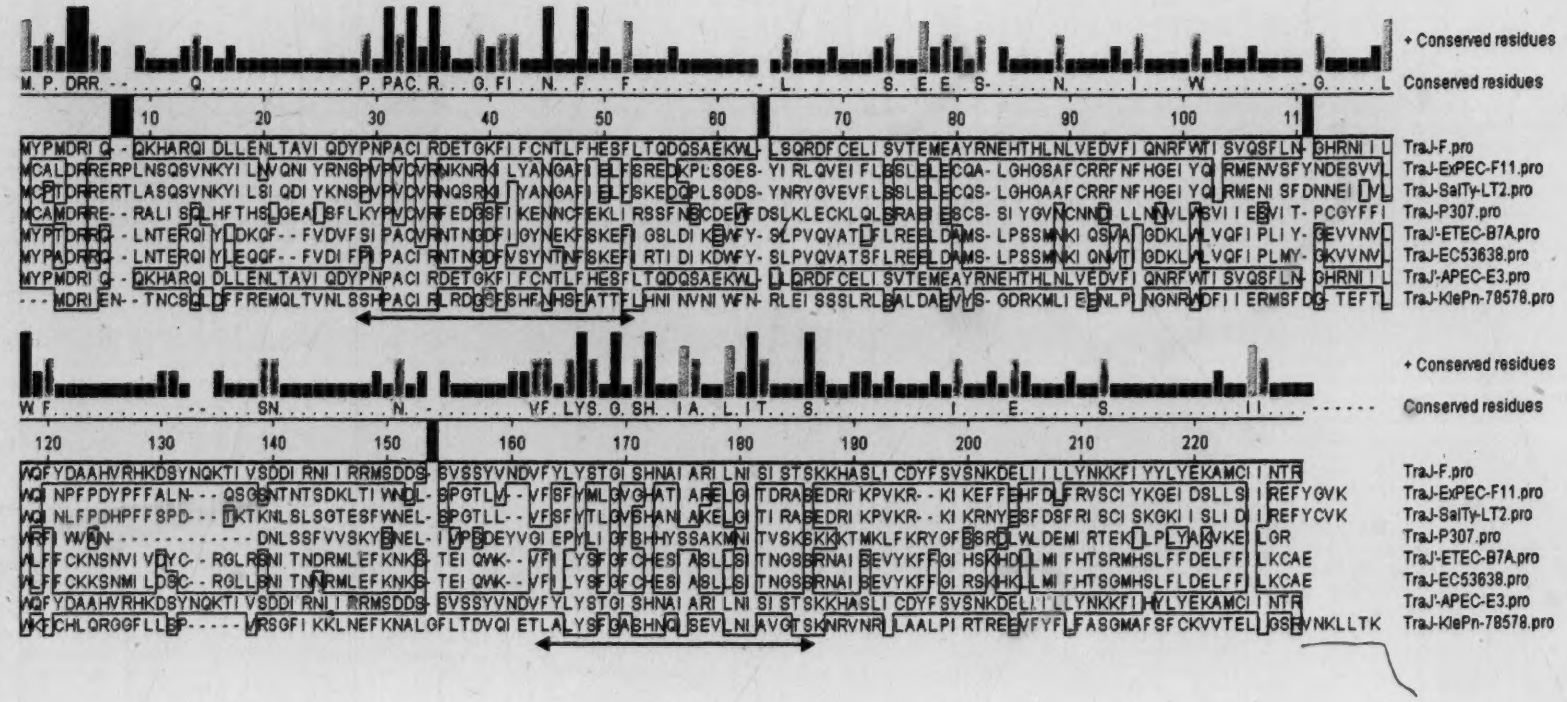

+ Conserved residues
Conserved residues
10
20
30
40
50
60
70
TraJ-F.pro
TraJ-R100.pro
TraJ-R1-19'.pro
78
75
79
+ Conserved residues
Conserved residues
80
90
100
110
120
130
140
150
TraJ-F.pro
TraJ-R100.pro
TraJ-R1-19'.pro
156
151
155
+ Conserved residues
Conserved residues
160
170
180
190
200
210
220
TraJ-F.pro
TraJ-R100.pro
TraJ-R1-19'.pro
229
223
222

Figure 3.2 Sequence alignment of F TraJ (Accession: BAA97942) and seven homologues from different strains and plasmids (Accession: ZP_03035127, NP_490561, AAA25523, ZP_03031175, YP_001919411, YP_001711968, YP_001338611). Residues that match F TraJ sequence are shown on top of the alignment as tall bars. A ruler is set in reference to F TraJ and grey areas refer to alignment adjustments in the F TraJ sequence. Boxed residues are conserved amino acids compared to the F TraJ sequence. Two clusters of conserved amino acids can be observed, one between residues 28 to 53 and between residues 162 to 187, as indicated by double-headed arrows.


+ Conserved residues
Conserved residues
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
TraJ-F.pro
TraJ-ExPEC-F11.pro
TraJ-SalTy-LT2.pro
TraJ-P307.pro
TraJ'-ETEC-B7A.pro
TraJ-EC53638.pro
TraJ'-APEC-E3.pro
TraJ-KlePn-78578.pro

II. Site-directed mutagenesis of specific residues in the putative HTH site in F TraJ.

Once a putative DNA-binding domain was suggested for F TraJ, mutagenic oligonucleotides were designed for conserved amino acids in this region (162-187). Mutations were introduced into pBAD24TraJ using the Quickchange procedure (Stratagene). Potential mutants were verified by DNA sequence analysis and protein production was verified by immunoblot analysis with anti-TraJ antiserum (Figure 3.3). These mutant proteins, produced *in trans*, were then tested for complementation of F*lac traJ90* (*traJQ26amber*), which gives the cell an F^- phenotype (Achtman *et al.*, 1971). The three *traJ* point mutants introduced into conserved residues were pJMR34 (G169R), pJMR56 (Y166D) and pJMR78 (H172D). The mutants were then inserted into MC4100/F*lac traJ90*, and mating ability was tested for each strain carrying one of the altered *traJ* proteins, as described in the methods section. The recipient strain was ED24 and the transconjugants were selected using minimal media with lactose and spectinomycin. The mating results are shown in Table 3.1. As hypothesized, the three point mutations reduced mating efficiency by approximately three orders of magnitude in comparison to wild type TraJ. Furthermore, the pJMR34 (G169R) mutant showed a slightly lower efficiency compared to the pJMR56 (Y166D) or the pJMR78 (H172D) mutants.

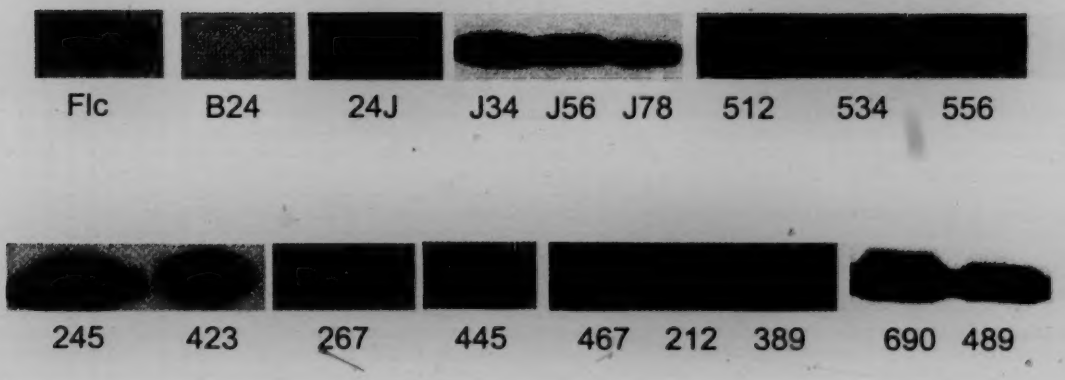
All three mutants not only changed the amino acid sequence of the protein, but probably also altered its secondary structure. Arginine is a basic amino acid that is larger than glycine; on the other hand, aspartic acid introduced an acidic residue in place of basic histidine or tyrosine. Therefore, new point mutants were made for the same TraJ residues in order to avoid these concerns. The new mutants are: pJMR512 (G169A),

Table 3.1 Mating assay results for the TraJ point mutants inside and outside the putative DNA-binding domain.

Donor strain MC4100/F*lac traJ90/*	**Mating efficiency*** (Transconjugants)/(100donors)
pBAD24TraJ (WT)	50
pBAD24 (vector control)	0.01
pJMR34 (G169R)	0.03
pJMR56 (Y166D)	0.1
pJMR78 (H172D)	0.05
pJMR512 (G169A)	0.2
pJMR534 (Y166A)	10
pJMR556 (H172A)	8.5
pJMR245 (D5A)	25
pJMR423 (Q14K)	71.4
pJMR267 (P31A)	75
pJMR445 (S65A)	66.7
pJMR467 (E77A)	75
pJMR212 (W118A)	57.1
pJMR389 (S186A)	33.3
pJMR690 (C224A)	50
pJMR489 (I225L)	50

*Average of two individual experiments

Figure 3.3 Immunoblots of the samples used in the mating assays for the TraJ point mutants expressed from pBAD24. Flc = WT F TraJ, B24 = negative plasmid control, 24J = positive plasmid control, J34-J56-J78 = Original 'HTH' point mutants, 512-534-556 = Alanine 'HTH' point mutants, 245-423-267-445-467-212-389-690-489 = JMR point mutants.


Flc
B24
24J
J34
J56
J78
512
534
556
245
423
267
445
467
212
389
690
489

pJMR534 (Y166A) and pJMR556 (H172A). These mutants were then tested in the same manner as the ones before and mating efficiency was determined, as shown in Table 3.1. The effect of the mutations was not as strong as with the previous mutants. However, mating efficiency was lower for all the mutants in comparison to wild type. Mating efficiency was slightly lower for both pJMR534 (Y166A) and pJMR556 (H172A), and two orders of magnitude lower for pJMR512 (G169A).

Several other point mutants were constructed and tested for activity in the same manner as the ones in the putative DNA-binding site. None showed significant difference in activity compared to wild-type TraJ (Table 3.1). These mutants include: pJMR245 (D5A), pJMR423 (Q14K), pJMR267 (P31A), pJMR445 (S65A), pJMR467 (E77A), pJMR212 (W118A), pJMR367 (I181A), pJMR389 (S186A), pJMR690 (C224A) and pJMR489 (I225L).

III. F TraJ C-terminal deletion mutagenesis.

Once the importance of the putative DNA-binding site was shown, more mutations were planned in order to determine if the non-conserved sequence distal to this motif is important for mating ability. Thus, deletion mutants were made for a shorter version of *traJ* in the pBAD24 vector, as explained in the methods section.

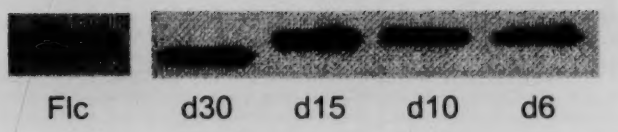
The first mutant tested was pJMRΔ30, which includes only the first 199 amino acids of TraJ. As shown in Table 3.2, conjugation was abolished in this mutant as well as in mutants with truncations at aa 214 (pJMRΔ15) and aa 219 (pJMRΔ10) (Table 3.2). Mutant pJMRΔ6, which was also not able to complement the *traJ90* mutation, was reconstructed by introducing an ochre mutation at aa 224 to give pJMR678 (*traJ224ochre*).

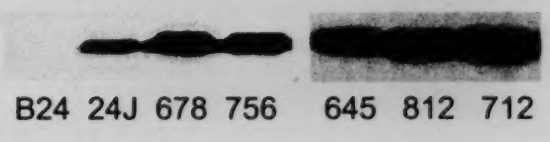
Table 3.2 Mating assay results for the TraJ C-terminus deletion mutants.

Donor strain MC4100/F*lac traJ90/...*	**Mating efficiency*** (Transconjugants)/(100xdonors)
pBAD24TraJ (WT)	50
pBAD24 (vector control)	0.01
pJMRΔ30 (TraJΔ30)	0.067
pJMRΔ15 (TraJΔ15)	0.05
pJMRΔ10 (TraJΔ10)	0.025
pJMRΔ6 (TraJΔ6)	0.1
pJMR678 (C224ochre)	0.064
pJMR756 (I225ochre)	0.2
pJMR812 (I226ochre)	0.077
pJMR645 (R229ochre)	32
pJMR712 (N227A)	33.3

*Average of two individual experiments

Figure 3.4 Immunoblots of the samples used in the mating assays for the TraJ deletion mutants expressed from pBAD24. Flc = WT F TraJ, B24 = negative plasmid control, 24J = positive plasmid control, d30- d15-d10-d6 = JMRΔ30-Δ15-Δ10-Δ6 deletion mutants. 678, 756, 645, 812, 712 = JMR ochre mutants.


Flc
d30
d15
d10
d6


B24
24J
678
756
645
812
712

This mutant was then tested in the same manner for mating efficiency, giving results that were similar to pJMRΔ6 (Table 3.2). To delineate the importance of the C-terminal tail, further mutants were made and tested. Mutants pJMR756 (*traJR225ochre*) and pJMR812 (*traJR226ochre*) were inactive, whereas mutant pJMR645 (*traJR229ochre*) complemented F*lac traJ90* as shown in Table 3.2. Interestingly, a point mutation in the tail, namely pJMR712 (N227A), showed no reduction in mating efficiency. As with previous results, mutant proteins were observed by western immunoblot to ensure their presence (Figure 3.4).

IV. Wild-type F TraJ *in vivo* DNA binding assays.

Due to the low solubility of purified TraJ, DNA-binding assays were performed *in vivo* using the chromatin immuno-precipitation (ChIP) assay for MC4100/*Flac* or MC4100/F*lac traJ90*. The procedure for the ChIP assay was detailed in the Methods section and involved cross-linking TraJ to the DNA with formaldehyde, sonication and immuno-precipitation of the complex with anti-TraJ antibodies. This was followed by heating to break the cross-links and PCR was performed with appropriate primers to detect the P_Y region (oligonucleotides RWI91-92). PCR gave a band on an agarose gel of the expected size for the sample with wild type *traJ*, whereas no band was observed for the sample with the *traJ90* amber mutant (Figure 3.6A). The expected DNA band is 200 bp long and includes the end of the *traJ* gene and the intergenic region between *traJ* and *traY* as well as the inverted repeat considered to be the site of TraJ binding (*sbj*) in R100 (shown in Figure 3.5). DNA was immunoprecipitated with anti-TraJ antiserum but not anti-TraK antiserum. Moreover, oligonucleotides for the F *traC* gene (traC-5 and traC-3), used as a negative control, did not show a band after amplification (Figure 3.6A).

Figure 3.5 Sequence of F *tra* DNA involved in the ChIP assay. Important sequences (primer sequences, stop and start codons and the inverted repeat) are shown in uppercase. Arrows starting in a circle represent oligonucleotides RWI91 and RWI92, used for PCR. Arrows starting with a diamond represent the stop and start codons for both *traJ* and *traY*, respectively. Inverted arrows represent the inverted repeat considered to be the site of binding for TraJ, *sbj*.

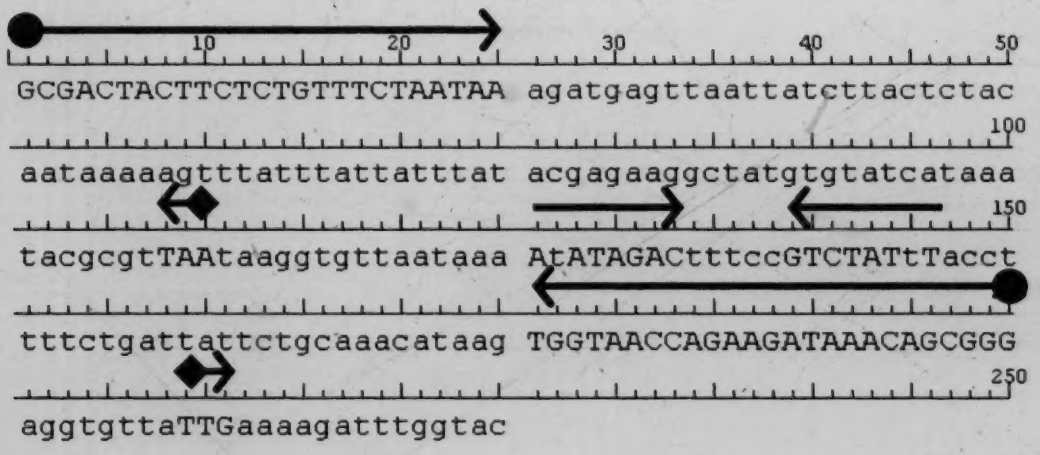
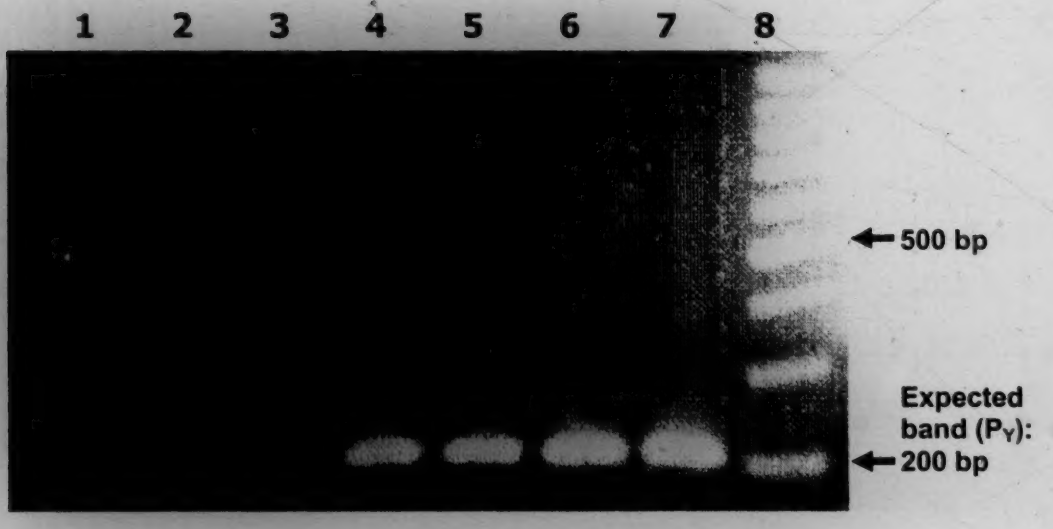
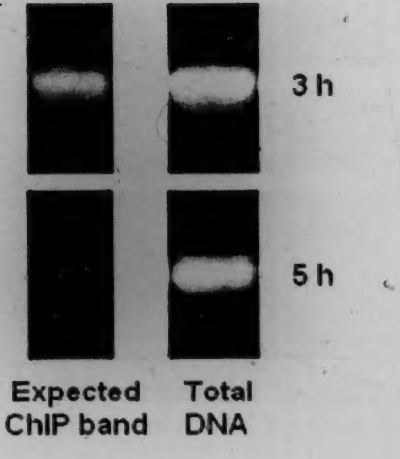

10
20
30
40
50
GCGACTACTTCTCTGTTTCTAATAA agatgagttaattatcttactctac
100
aataaaaagtttatttattatttat acgagaaggctatgtgtatcataaa
150
tacgcgtTAAtaaggtgttaataaa AtATAGACtttccGTCTATtTacct
tttctgattattctgcaaacataag TGGTAACCAGAAGATAAACAGCGGG
250
aggtgttaTTGaaaagatttggtac

Figure 3.6 PCR results in duplicate from the ChIP assay of TraJ run on a 1.5% agarose gel. A) Results for samples, obtained after 3 hour growth: MC4100/F*lac* (anti-TraK antibodies) (1), MC4100/F*lac* (traC-5,-3 primers) (2), MC4100/F*lac traJ90* (3), MC4100/F*lac* (4), MC4100/F*lac* (duplicate) (5), total DNA for MC4100/F*lac* (6) and MC4100/F*lac traJ90* (7), 1 kb DNA ladder (8). B) Results for samples taken at 3 hours (exponential phase) and 5 hours (stationary phase), the band shown is 200 kb in size.

A)


1
2
3
4
5
6
7
8
500 bp
Expected band (P_Y):
200 bp

B)


3 h
5 h
Expected ChIP band
Total DNA

TraJ has been shown to be present in stationary phase; however conjugative transfer is not active. Therefore, TraJ binding to DNA was also tested in late exponential phase and early stationary phase to see if DNA binding was reduced. Using the ChIP assay, it was shown that the band obtained from the sample taken at five hours was significantly lower than the one at three hours, while total DNA remain unchanged (Figure 3.6B).

C. Discussion

F TraJ contains a putative helix-turn-helix motif at position 162-187.

As shown by the alignment with its closest homologues in F-like plasmids, an HTH site was predicted in residues 162-187 of F TraJ. HTH motifs are 18 to 25 residues long and usually include a glycine at the turn (Pabo & Sauer, 1984). TraJ has these characteristics including a glycine at position 169. Therefore TraJ could bind DNA as shown for R100 TraJ (Taki *et al.*, 1998). Furthermore, TraJ from other F-like plasmids could also function in a similar manner. Although there is little sequence identity, TraJ from F-like plasmids could maintain structural similarity and contain the same DNA-binding domain.

Conserved residues close to the turn in the putative HTH motif are essential for TraJ activity.

Site-directed mutagenesis of three amino acids (G169, Y166 and H172) in the DNA-binding motif resulted in reduced mating ability in complementation assays. These three mutants were located within or close to the 'turn' of the HTH motif, which is usually found in the first half of the motif. Among most HTH bearing proteins, a glycine around position nine ('turn') of the motif (X^{1}-...-A(S)$^{\sim 5}$-...-G$^{\sim 9}$-...-I$^{\sim 15}$-...-X$^{\sim 20}$), has been

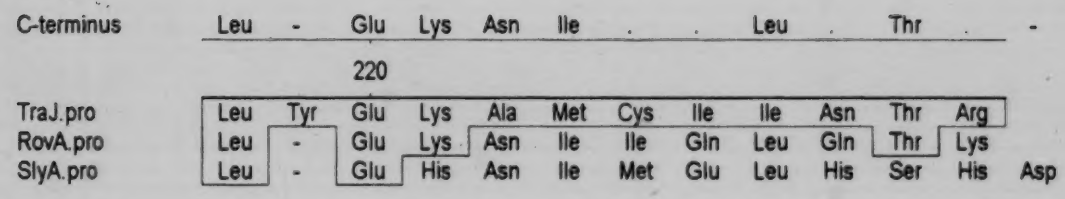
shown to be present (Pabo & Sauer, 1984). Both pJMR34 (G169R) and pJMR512 (G169A) confirmed the importance of this residue. However, some mutations in the glycine have no effect on activity (Pabo & Sauer, 1984), suggesting a certain degree of flexibility in the structure of this domain. For F TraJ, this flexibility was not found since both mutations at G169 reduced activity by over two orders of magnitude. Moreover, mutations both before and immediately after the turn (Y169 and H172) showed reduced mating ability, albeit not as strongly as in the case of the G169 mutations. Mutant plasmids pJMR56 (Y166D) and pJMR78 (H172D), showed over two orders of magnitude reduction in activity; however mutants pJMR534 (Y166A) and pJMR556 (H172A), showed a reduction of approximately ten-fold compared to wild-type TraJ. Mutations upstream or downstream of the motif showed no effect on mating ability. Mutations within or very close to the HTH site, namely pJMR367 (I181A) and pJMR389 (S186A), showed no effect on activity. This would suggest that there is some flexibility for the sequence of the helices of F TraJ motif, albeit not for the turn. Thus, the right angle required for DNA-binding was disrupted when the G169 at the 'turn' was mutated.

The last four residues of TraJ are essential for activation of mating ability in the F plasmid.

Deletion of the last four amino acids at the C-terminus of TraJ showed an F$^-$ phenotype in MC4100/F*lac traJ90*. These mutants were originally produced to test whether the sequence downstream of the putative HTH motif was essential. Deletion of the last thirty amino acids, pJMRΔ30, showed a strong decrease in mating ability whereas the level of the protein was unaffected. Shorter deletion mutants, pJMRΔ15 and pJMRΔ10, showed the same effect (Table 3.2). The sequence downstream of the stop codon is not needed for proper activity of the protein, as the pJMRΔ6

mutant was produced both by deleting the whole C-terminal region and by inserting an ochre mutation (*C224ochre*) by site-directed mutagenesis (pJMR678). The latter mutant failed to complement F*lac traJ90* as did an ochre mutation at I225 (pJMR756) and I226 (pJMR812). Mating ability regained almost wild-type levels for the ochre mutation in R229 (pJMR645), suggesting the terminal four amino acids were essential for function. Moreover, a single amino acid substitution within the last four amino acids, namely N227A (pJMR712), did not show any effect on mating ability, suggesting that the importance of the C-terminus involves its length, not its sequence. The mechanism by which this domain works is still to be determined, but some clues have been found. The activator RovA also has an essential C-terminal tail of four amino acids that is similar to TraJ in amino acid charges (Figure 3.7). This effect was shown to be length-dependent and required last four amino acids, as gradual complementation of activity was seen for shorter (-3, -2 and -1) deletion mutants (Tran *et al.*, 2005). RovA, which is the main activator of the *inv* genes in *Yersinia* and part of the SlyA/MarR family, has been shown to both counteract H-NS silencing and also interact directly with RNA polymerase to activate the *inv* promoter (Tran *et al.*, 2005). RovA, like other members of the SlyA/MarR family, binds DNA via a winged-helix domain (Ellison & Miller, 2006). The winged-helix domain of RovA involves an HTH motif followed by two β-strands (Ellison & Miller, 2006). F TraJ has been shown to be important for disrupting H-NS silencing at the P_Y promoter, possibly by a mechanism similar to RovA (this work). This suggests that TraJ and RovA probably share some structural homology. Since the C-terminal domain of RovA is thought to interact with the RNA polymerase (Tran *et al.*, 2005), TraJ could also bind RNA polymerase via its C-terminus, as an alternative sigma factor. In addition, the C-terminal region of TraJ could be part of the double β-strand (after

Figure 3.7 Sequence comparison between the C-terminal regions of F TraJ (Accession: BAA97942), RovA (Accession: CAH21526) and SlyA (Accession: AAL20366). The row on the top is the consensus amino acid sequence for the three proteins. The ruler is set to F TraJ sequence.


C-terminus
Leu - Glu Lys Asn Ile . . Leu . Thr . -
220
TraJ.pro Leu Tyr Glu Lys Ala Met Cys Ile Ile Asn Thr Arg
RovA.pro Leu - Glu Lys Asn Ile Ile Gln Leu Gln Thr Lys
SlyA.pro Leu - Glu His Asn Ile Met Glu Leu His Ser His Asp

HTH) arrangement of a RovA-like winged-helix domain. Thus, considering this two possibilities, TraJ could be a member of the MarR/SlyA family of regulators. In order to demonstrate this, the crystal structure of TraJ would be needed.

Wild-type F TraJ binds to the *traYX* promoter in vivo.

Using the ChIP assay, it was shown that TraJ binds *in vivo* to a 200 bp fragment of DNA that includes the last 110 bp of *traJ* and 90 bp of the intergenic region between *traJ* and *traY*. The intergenic region includes both the -35 and -70 boxes of P_Y, as well as the inverted repeat sequence that functions as the site of binding of TraJ, *sbj*, in R100 (Taki *et al.*, 1998). This indicates that F TraJ probably activates the *traYX* operon by binding to its promoter region. Although, these results do not indicate precisely where it binds nor its mechanism of action, they show that TraJ has high affinity for this region and could act as a countersilencer as suggested previously (Will & Frost, 2006a). Furthermore, the decrease in intensity of the DNA band over time could indicate that H-NS competes with TraJ, leading to silencing and the F^- phenotype in stationary phase. However, the TraJ protein could reactivate the system once environmental and nutritional conditions improve (Frost & Manchak, 1998), most probably by competing with H-NS for the *traY* promoter.

Chapter 4. F TraJ is a 226 amino acid protein that forms dimers and higher-order complexes

A. Introduction

F plasmid TraJ is considered to be a 229 amino acid protein that activates the promoter of the main transfer operon (Frost *et al.*, 1994). TraJ activates conjugation by binding to the *traY* promoter region and impeding repression by the host factor H-NS (this work)(Will & Frost, 2006a). However, TraJ activity is limited to growing cells, as cells acquire an F^- phenotype once they enter stationary phase (Frost & Manchak, 1998). This phenotype is attributed to H-NS silencing when conditions are disadvantageous for the cell (Will *et al.*, 2004). Why TraJ does not counteract H-NS in stationary phase remains unknown, however post-translational changes to TraJ have been suggested (Lau-Wong *et al.*, 2008).

Although, TraJ is inessential to activate the *traY* promoter when cells lack H-NS (PD32 (*hns*)/Flac *traJ90*), there is hyper activation of P_Y if TraJ is expressed in a *hns* knockout (Will & Frost, 2006a). This effect was considered to be due to up-regulation of H-NS paralogues such as StpA (Will & Frost, 2006a). However, the possibility that TraJ promotes transcription cannot be disregarded. This putative dual desilencer-activator characteristic, along with the requirement for the C-terminal region of TraJ, suggests that it could be related to RovA from the MarR/SlyA family of activators (Tran *et al.*, 2005). Regulators of this family act by binding DNA, generally as homodimers. Although TraJ dimers have not been demonstrated, it has been suggested previously since R100 TraJ binds to a palindromic inverted repeat in the sequence upstream of its cognate *traY* gene (Taki *et al.*, 1998).

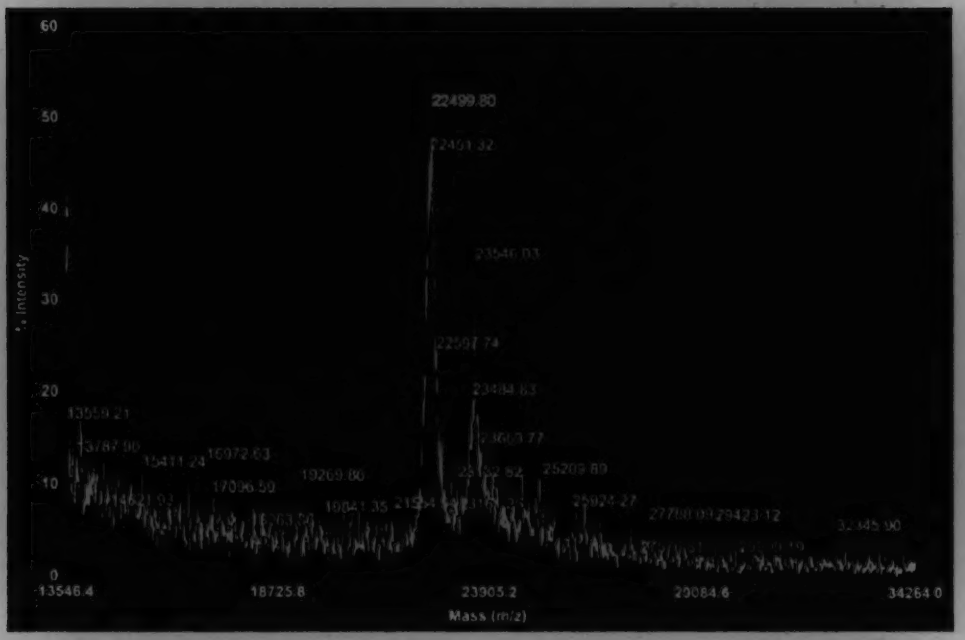
The objective of this study was to identify the mass of TraJ in both exponential and early stationary phase by mass spectrometry (MS) to detect modifications. An unexpected result from MS suggested that the fourth methionine rather than the first methionine was the true start codon for TraJ. In addition, the oligomeric state of TraJ was tested both in the presence and absence of H-NS, as well as in a F*lac traJ90* background.

B. Results

I. Mass Spectrometry data for F TraJ.

Two samples from a culture of MC4100/F*lac traJ90*/pBAD24TraJ were isolated at three hours and at five hours of growth, which represented exponential phase and early stationary phase, respectively. Since TraJ is very difficult to purify and it loses activity rapidly, protein samples for MS were isolated from a polyacrylamide gel of a crude preparation of TraJ. Samples were electrophoresed in duplicate with sterile equipment and filtered solutions on a 12% polyacrylamide resolving gel, and the band representing overexpressed TraJ was excised with a sterile scalpel. MC4100/F*lac traJ90*/pBAD24 was used as a negative control. Samples were placed in ddH_2O and sent for mass spectrometric analysis. The results were not completely satisfactory, as the sample at five hours showed a high background and no specific peak (not shown). On the other hand, the sample at three hours showed a particularly strong peak of approximately 50% signal intensity, although it did not match the predicted mass for TraJ of 27.06 kDa. This peak corresponded to a mass of 22.49 kDa as shown in Figure 4.1. Identical results were obtained for duplicate samples. These results suggested that TraJ itself was not detectable whereas a fragment resulting from post translational modification or proteolytic digestion was present.

Figure 4.1 Mass spectrometry results for the TraJ sample obtained from the strain MC4100/F*lac traJ90*/pBAD24TraJ at three hours of growth. The number 22499.80 represents the peak that most closely resembles the expected size of TraJ. The X-axis represents the mass-to-charge ratio, whereas the Y-axis is the percentage of signal intensity.


60
50
40
30
20
10
0
% Intensity
22499.80
24491.32
23546.03
22597.74
23484.63
13559.21
13787.90
15411.24
16072.63
17006.59
19269.80
25209.89
25924.27
27780.09
29423.12
32445.90
13546.4
18725.8
23905.2
29084.6
34264.0
Mass (m/z)

II. Translational initiation of F TraJ.

Using a program from the ExPASy proteomics server (Swiss Institute of Bioinformatics), possible fragments of TraJ that would lead to a 22.499 kDa mass were suggested (not shown). Out of the five most likely candidates, one was considered to be the most interesting. This 22.496 kDa polypeptide corresponded to TraJ lacking the last 30 amino acids as well as the first three aa. Thus, it would start at the codon at position four (M4) instead of position one (M1). The M4 position is preceded by a possible ribosome binding site (RBS), further suggesting it might be the correct start site. This possibility was further characterized by deleting the first three codons from *traJ* in pBAD24TraJ, giving pJMRΔ3N, which was tested for activity and expression. pJMRΔ3N activated conjugation in MC4100/F*lac traJ90* and it was detectable by immunoblot, suggesting that the protein was present and functional. However, the P_{BAD} promoter in this construct might produce the protein from another start codon within the multiple cloning site. To check this, mutagenic oligonucleotides were designed in order to mutate both methionine codons, M1 and M4, to threonine codons. These mutations (M1T and M4T) were introduced into pILJ11 which is the pBCSK$^+$ vector with an F SalI-PstI insert containing the whole *traJ* gene along with part of *traM* and *traY*, including both intergenic sequences upstream and downstream of the *traJ* gene (Table 2.2). Three mutants, pJM1T, pJM4T and pJM1-4T, were assayed for activity and protein expression. As hypothesized, the mutant at the first methionine (pJM1T) produced TraJ at normal levels and complemented the *traJ90* mutation. On the other hand, the mutant at position four (pJM4T) failed to produce the protein and thus, was incapable of activating conjugation. Finally the mutant in both M1 and M4 (pJM1-4T) showed a phenotype identical to the pJM4T mutant in that it did not complement the *traJ90* mutation and did not produce protein (Figure 4.2). Thus the methionine at

Figure 4.2 Immunoblots of the TraJ methionine mutants expressed from pBCSK$^+$ with controls. BCSK = negative control, J11 = positive control, M14 = double methionine mutant, M1T = mutant in first methionine, M4T = mutant in fourth methionine

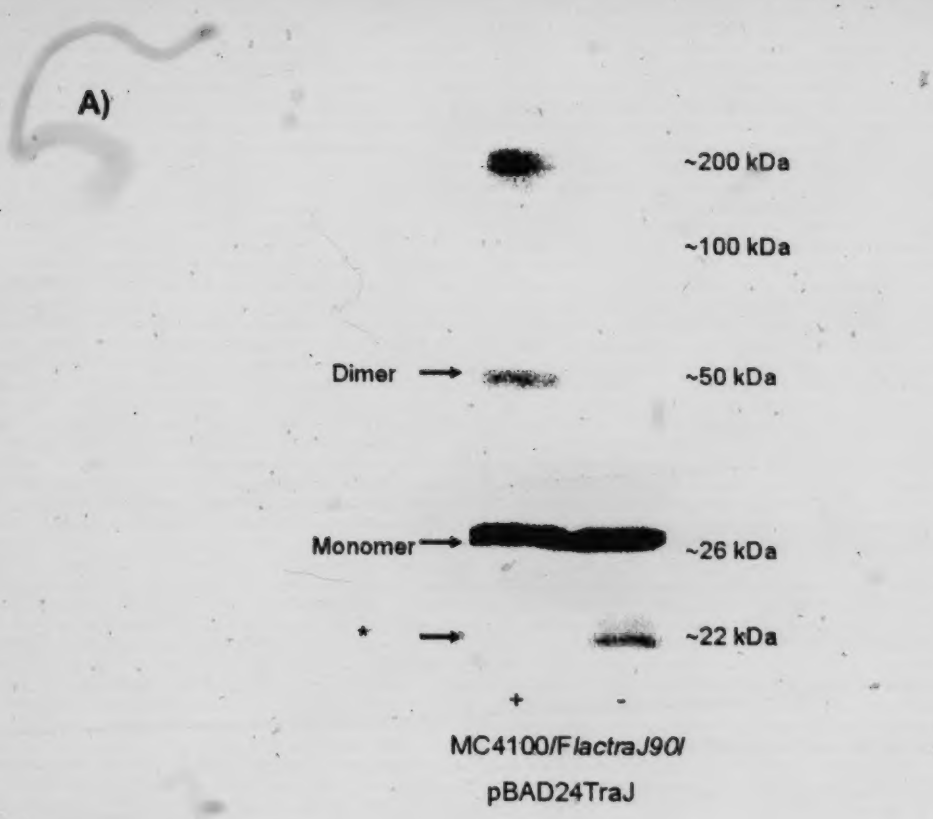
BCSK J11 M14 M1T M4T

position four is probably the correct start codon for *traJ* and the size of TraJ should be revised to 226 amino acids.

III. TraJ oligomerization in an F^+ and *hns*⁻ strain.

In order to test the oligomeric state of TraJ, cross-linking experiments were performed for MC4100/F*lac traJ90*/pBAD24TraJ. The strain was grown in LB with 0.1% arabinose to late exponential phase (OD_{600} ~ 1.0). Cross-linking was performed with DSS and quenched with Tris buffer, whereas a sample with no DSS was used as negative control. Immunoblot analysis showed a band of the right size for TraJ (~ 27 kDa) and also a band of double this size (~ 54 kDa) which suggests dimer formation only in the cross-linked sample (Figure 4.3). In order to determine whether the C-terminal region is important for oligomerization, MC4100/F*lac traJ90*, carrying either the plasmid pJMRΔ6 or pJMRΔ30, was crosslinked with DSS in the same manner (Figure 4.3B). Results for both proteins were similar to the wild-type, although both bands, representing the monomer and the dimer, had a greater mobility in the gel, as expected because TraJ was truncated. Finally, to determine whether the dimers were due to TraJ-H-NS interactions, the strains PD32/F*lac* (Figure 4.4A) or F*lac traJ90* (MC4100 *hns* mutant, Ap^r) along with either pBAD33 (Cm^r), pIJL14 (pBAD33TraJ) or pIJL14Δ6 were tested (not shown). Surprisingly, the immunoblot for PD32/F*lac* producing wild-type TraJ consistently showed multiple bands and smearing with no single bands being observed (Figure 4.4A). Cross-linking was performed in the absence of F*lac*, using strains PD32/pIJL14 and PD32/pIJL14Δ6 (Figure 4.4B). TraJ from these plasmids formed monomers and dimers of the predicted sizes as observed in Figure 4.4.

Figure 4.3 Cross-linking with DSS for MC4100/F*lac traJ90* samples with either wild-type TraJ (A) or truncated (Δ6 or Δ30) versions of TraJ (B). On the right side are the approximate masses for reference proteins. The symbol '+' means the cross-linker was added to the sample, whereas the symbol '-' means there was no cross-linker added. The asterisk (*) indicates the commonly observed degradation product that might correspond to the 22.499 peak seen by MS.


A)
~200 kDa
~100 kDa
Dimer
~50 kDa
Monomer
~26 kDa
*
~22 kDa
+
-
MC4100/F*lactraJ90I*
pBAD24TraJ

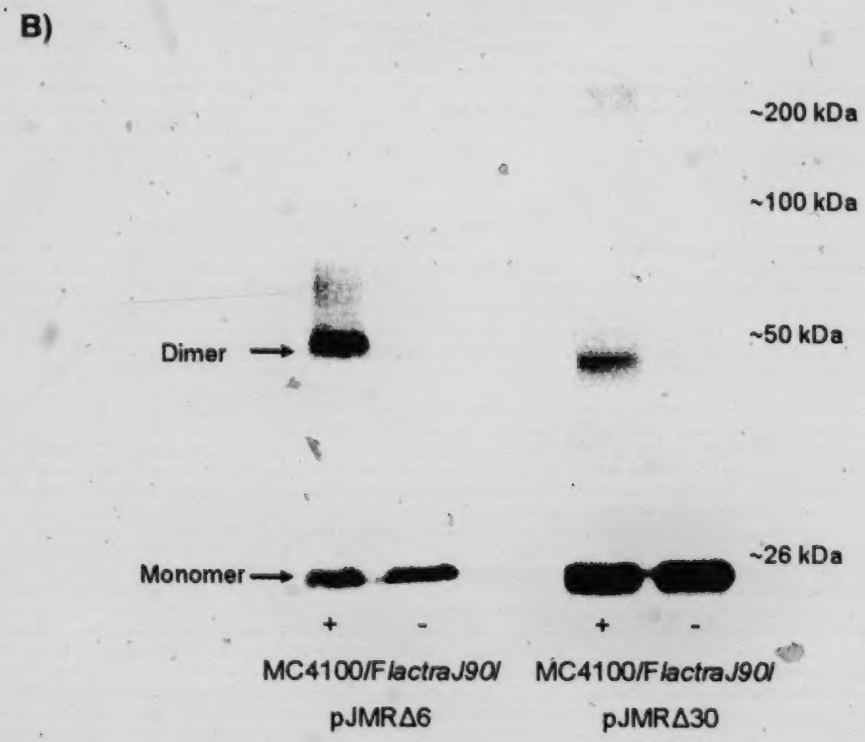
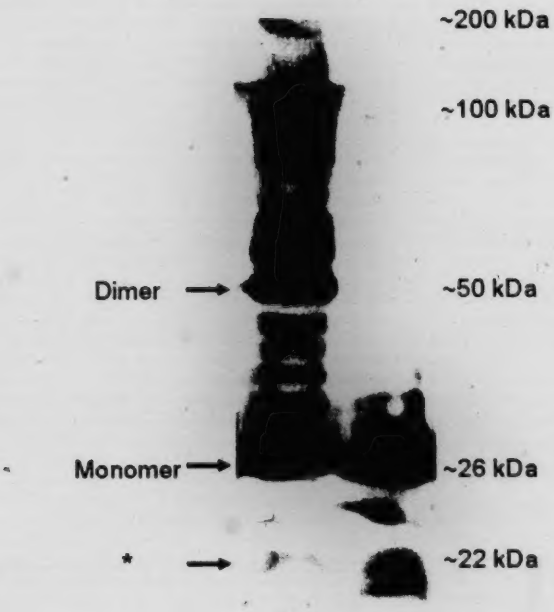

B)
~200 kDa
~100 kDa
Dimer
~50 kDa
Monomer
~26 kDa
+
-
+
-
MC4100/F*lactraJ90I*
pJMRΔ6
MC4100/F*lactraJ90I*
pJMRΔ30

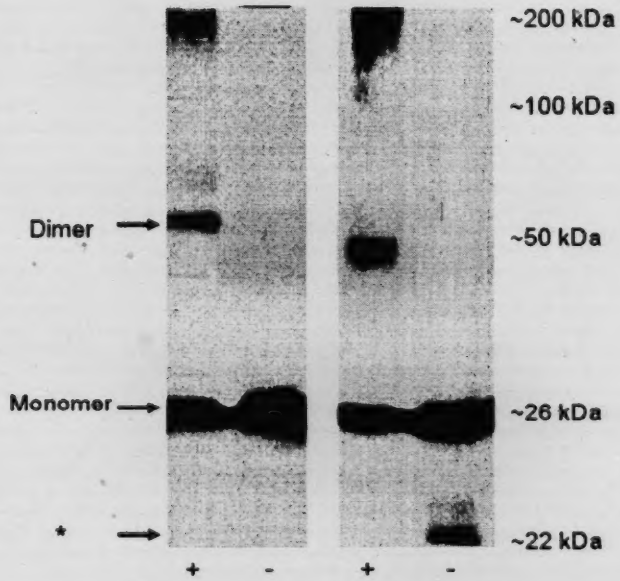
Figure 4.4 Cross-linking with DSS for PD32 (*hns*) samples with either wild-type TraJ from F*lac* (A) or TraJ (wild-type or Δ6) expressed from the pBAD33 vector with no F present (B). On the side are specified the approximate kDa values for protein standards. The symbol '+' means the cross-linker was added to the sample, whereas the symbol '-' means there was no cross-linker involved. The asterisk (*) indicates the degradation product of TraJ.

A)



PD32/F*lac*

B)



PD32/pILJ14 PD32/pILJ14Δ6

C. Discussion

Characterization of a proteolytic fragment of TraJ of ~22.5 kDa.

Results from MS suggested that there is a form of TraJ which is smaller than the original protein. This form is stable and seems to be the result of specific degradation that occurs as the cell goes into late exponential phase (Isabella Lau, unpublished results). There are two other lines of evidence that would support this claim. First, TraJ is present but inactive in late stationary phase, suggesting that the protein is inactivated by either protein modification or the lack of a factor required for its function. Second, a band of approximately 22.5 kDa was routinely detected in many different TraJ immunoblots (see Figure 4.3A for an example). This truncated version of TraJ appears to be a stable degradation product with an unknown function. Whole TraJ is difficult to be detected by MS under standard conditions, most probably because of its insolubility and resistance to "flying" (since time-of-flight is measured). Therefore, modified procedures might be useful to analyze this protein such as using a different matrix to support the sample. However, we did not pursue this since obtaining enough pure TraJ for these studies was not feasible at the time.

TraJ is a 226 amino acid protein with a start codon in Met4.

F TraJ, previously considered to be 229 amino acids long, has been shown here to be three residues shorter. The results from MS suggested a smaller protein that would start from the methionine at position four. This proved to be true, since a TraJ mutation in methionine one was still functional, whereas a mutant in the fourth methionine was not. Amino acid positions in this work are kept with the former numbering system for the sake of consistency with Frost *et. al.* (1994); however, they must be adjusted hereafter to the new start site with the actual protein being 226

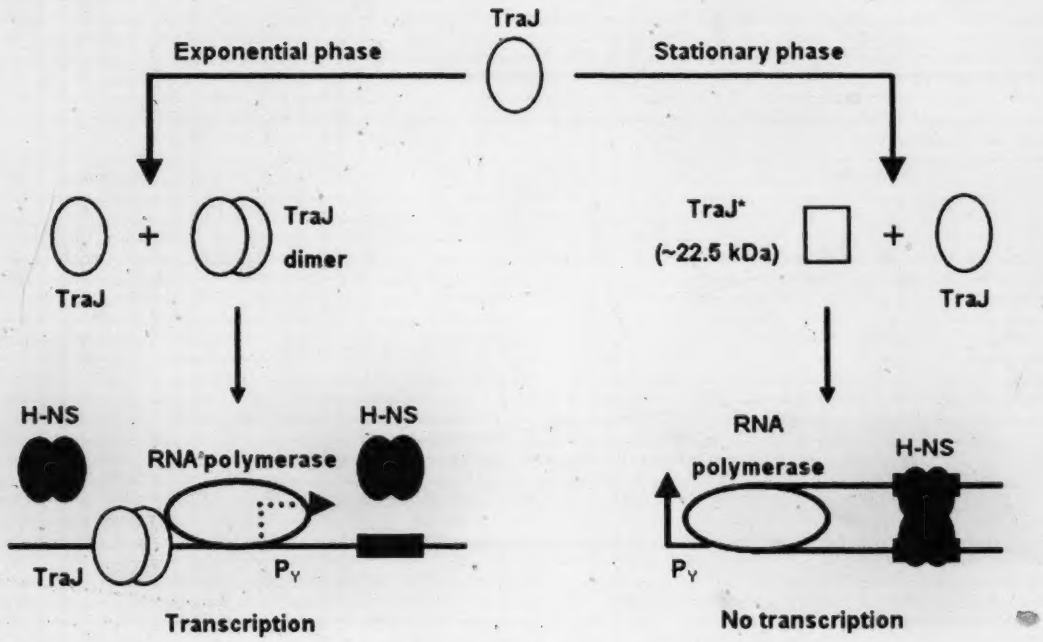
residues in length. Surprisingly, TraJ from R100, although it is highly conserved in this region, lacks M4 and probably uses the methionine equivalent to M1. However, R100 TraJ is shorter than F TraJ, possessing only 223 amino acids.

F TraJ can form homodimers *in vivo*.

Crosslinking with DSS demonstrated that TraJ is capable of forming homodimers. Most regulators act as dimers or even higher-order oligomers; for instance, F TraM regulates its own expression as a tetramer (Lu *et al.*, 2006) and F TraY as a dimer of a fused dimer (functional tetramer) form (Nelson & Matson, 1996). Moreover, regulators from the MarR/SlyA family are active as dimers in order to achieve their function. However, it has been shown that only 'fresh' TraJ is functional (Taki *et al.*, 1998). Therefore, the dimer conformation for TraJ might be the active form in the cell, and impeding oligomerization might be a method of inactivating TraJ in stationary phase (Figure 4.5). Surprisingly, in most of the samples that did not have cross-linker added, the 22.5 kDa band was noticeable on the immunoblot, whereas it was absent in the samples treated with DSS. This suggested that the 22.5 kDa polypeptide was capable of cross-linking to other proteins although these products could not be detected on the gel.

Samples from strains that lacked H-NS showed formation of several large complexes, albeit only in the presence of the F plasmid. On the other hand, having F in the background made almost no difference for samples obtained from an *hns*$^+$ strain. Since the three main F promoters, P_M, P_J and P_Y are derepressed in an *hns* mutant, this de-silencing could lead to overproduction of the proteins and to formation of nonspecific hetero-oligomeric complexes. TraJ could also form higher-order homo-oligomers (tetramers, hexamers, octamers); however the importance of these

Figure 4.5 Model for the TraJ oligomeric state in both exponential and stationary phase. TraJ can form dimers during exponential phase, albeit most of TraJ remains as a monomer. Dimer TraJ could activate RNA polymerase transcription. In stationary phase, TraJ is unable to form dimers and is then digested or modified to a stable but inactive form. Neither monomeric TraJ nor the modified or digested TraJ are able to alleviate H-NS silencing of P_Y.


TraJ
Exponential phase
Stationary phase
TraJ
+
TraJ
dimer
TraJ*
(~22.5 kDa)
+
TraJ
H-NS
H-NS
RNA polymerase
TraJ
P_Y
Transcription
RNA
polymerase
H-NS
P_Y
No transcription

complexes appears to be negligible as they only appear when H-NS is absent. Recently, aggregation of TraJ has been shown to be prevented by the addition of DTT (dithiothreitol), suggesting that the five cysteines in TraJ cause multimerization via intermolecular disulfide bond formation. The importance of a high number of cysteines and the formation of multiple disulfide bonds has already been proven for other F transfer proteins (Klimke *et al.*, 2005).

Chapter 5. Discussion and future directions

A. Discussion

The F plasmid regulatory system has been shown to be an intricate web of host- and plasmid-encoded factors. However, not only is the system complex, the members of this system are also complex in themselves. TraJ and H-NS have been suggested to be the two main players in the up- and down-regulation of the F conjugative T4SS, and their antagonism has been shown to be important (Will & Frost, 2006a). However, the mechanism of TraJ activity has been difficult to elucidate. A series of mutants were made throughout the TraJ sequence in order to shed light on its activity and mechanism of action (Figure 5.1).

F TraJ has been demonstrated to be a 226 (starting at M4), not 229 (starting at M1), amino acid protein that was proposed to bind to the P_Y promoter region, by means of its predicted HTH motif, in order to activate transcription of the F *traYX* operon. The binding site is probably the inverted repeat within the region amplified in the ChIP assay. This would agree with results for TraJ from the F-like plasmid R100, which is the closest relative to F TraJ (Taki *et al.*, 1998). Binding would then be essential for the disruption of the silencing complex formed by H-NS, either directly by competition for the binding site or by complexing to another factor, such as the RNA polymerase. Dimers of TraJ have also been found to form in exponential phase, although there is the possibility that the monomer form is also important. TraJ has been shown to be present throughout stationary phase; however this protein is inactive (Frost & Manchak, 1998). TraJ could bind DNA in the dimer form, which would be found in early exponential phase. Dimerization could then be impeded by modification or partial digestion as cells enter stationary phase (Figure 5.2). This modification of the protein may be reversible as the system is

Figure 5.1 Summary of most of the mutants used in this study. In the box is the amino acid that was changed for the one underneath. Bars on the top represent the phenotype of each specific mutation: a tall bar means very low or no activity (reduction in activity ≥ two orders of magnitude), a medium bar means low to medium activity (reduction in activity ~ one order of magnitude), a short bar means wild-type activity (no reduction in activity). The residues positions agree with the new size (226 amino acids) of the protein, thus M1 represents former M4.

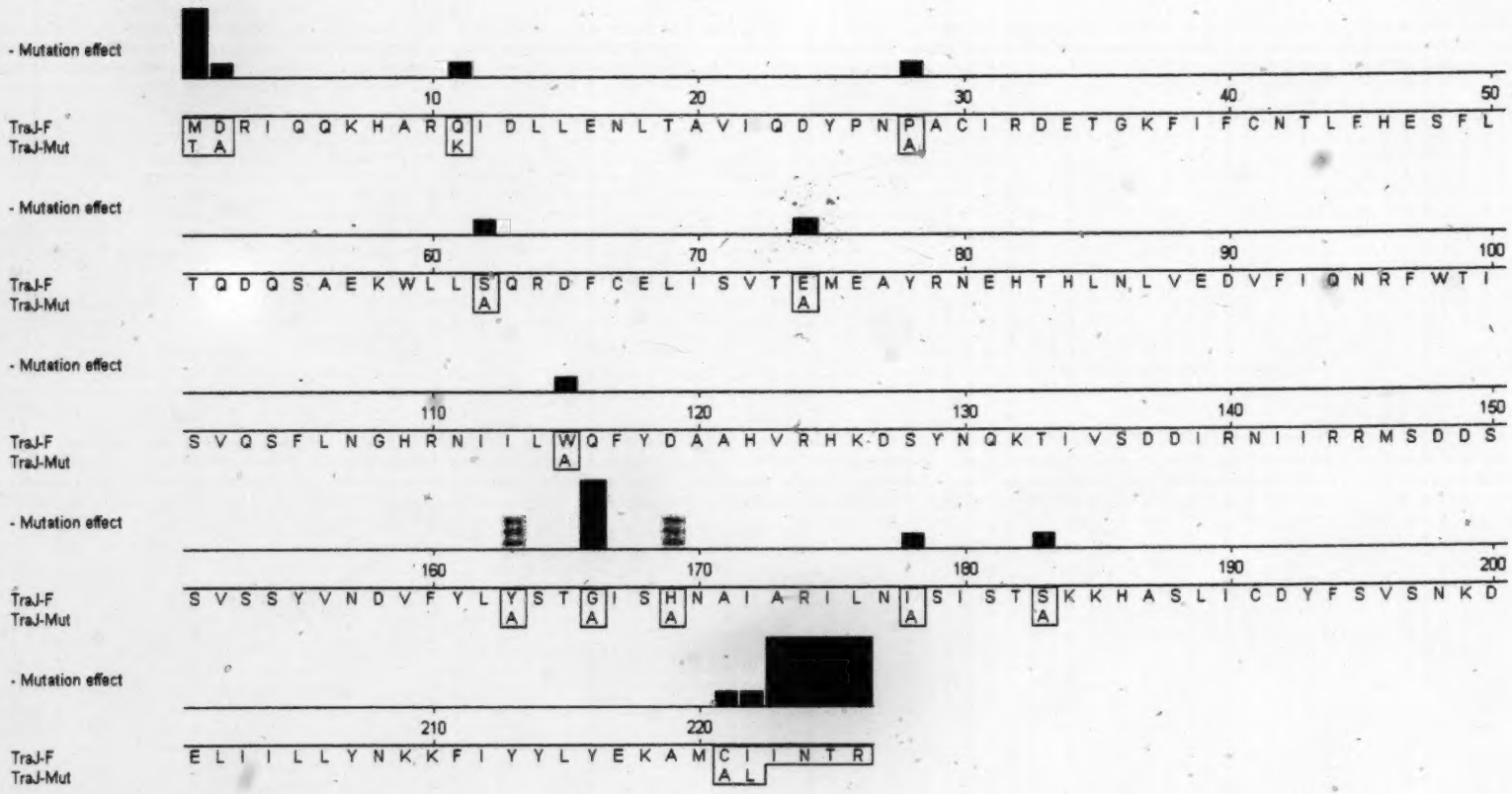
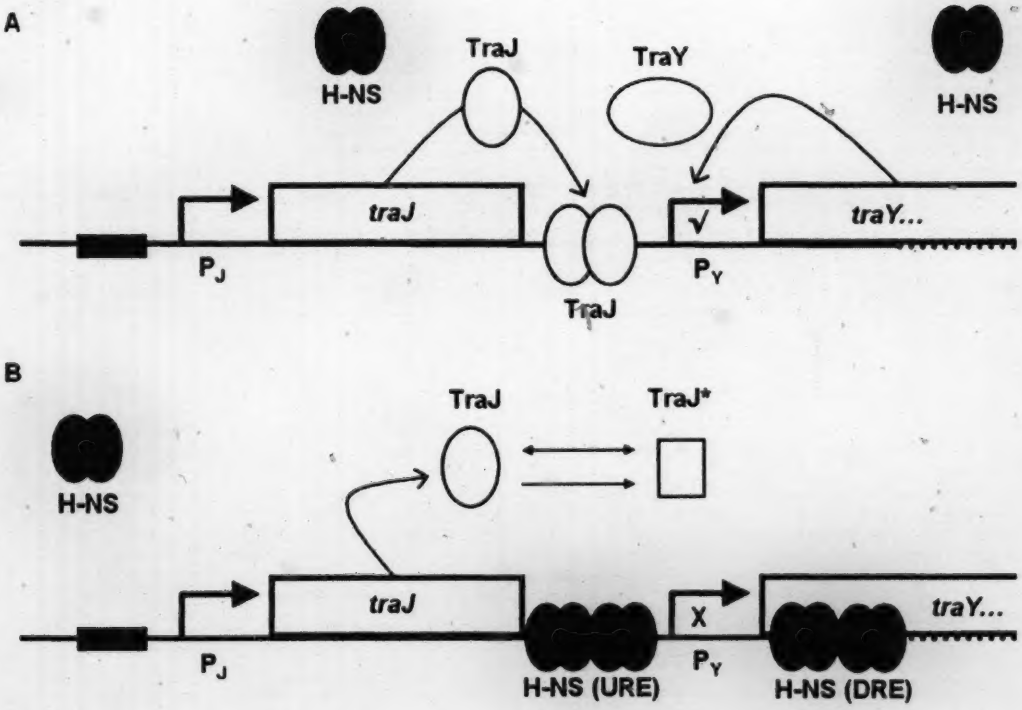

- Mutation effect
10
20
30
40
50
TraJ-F
TraJ-Mut
M D R I Q Q K H A R Q I D L L E N L T A V I Q D Y P N P A C I R D E T G K F I F C N T L E H E S F L
T A K A
- Mutation effect
60
70
80
90
100
TraJ-F
TraJ-Mut
T Q D Q S A E K W L L S Q R D F C E L I S V T E M E A Y R N E H T H L N L V E D V F I Q N R F W T I
A A
- Mutation effect
110
120
130
140
150
TraJ-F
TraJ-Mut
S V Q S F L N G H R N I I L W Q F Y D A A H V R H K D S Y N Q K T I V S D D I R N I I R R M S D D S
A
- Mutation effect
160
170
180
190
200
TraJ-F
TraJ-Mut
S V S S Y V N D V F Y L Y S T G I S H N A I A R I L N I S I S T S K K H A S L I C D Y F S V S N K D
A A A A A A
- Mutation effect
210
220
TraJ-F
TraJ-Mut
E L I I L L Y N K K F I Y Y L Y E K A M C I I N T R
A L

Figure 5.2 Model for TraJ regulation mechanism. A) Early exponential phase activation of the *tra* operon by binding of a TraJ dimer to *sbj*. The TraY protein, once produced, further induces its own promoter as well as P_M. B) During early stationary phase, silencing of the *tra* operon occurs by H-NS, which binds to the upstream and downstream regulatory elements (URE, DRE) of P_Y. TraJ's ability to form dimers is reduced and the TraJ monomer is modified and/or partially digested (TraJ*), thus rendering the protein inactive.


A
H-NS
TraJ
TraY
H-NS
traJ
traY...
P_J
P_Y
TraJ
B
TraJ
TraJ*
H-NS
traJ
X
traY...
P_J
H-NS (URE)
P_Y
H-NS (DRE)

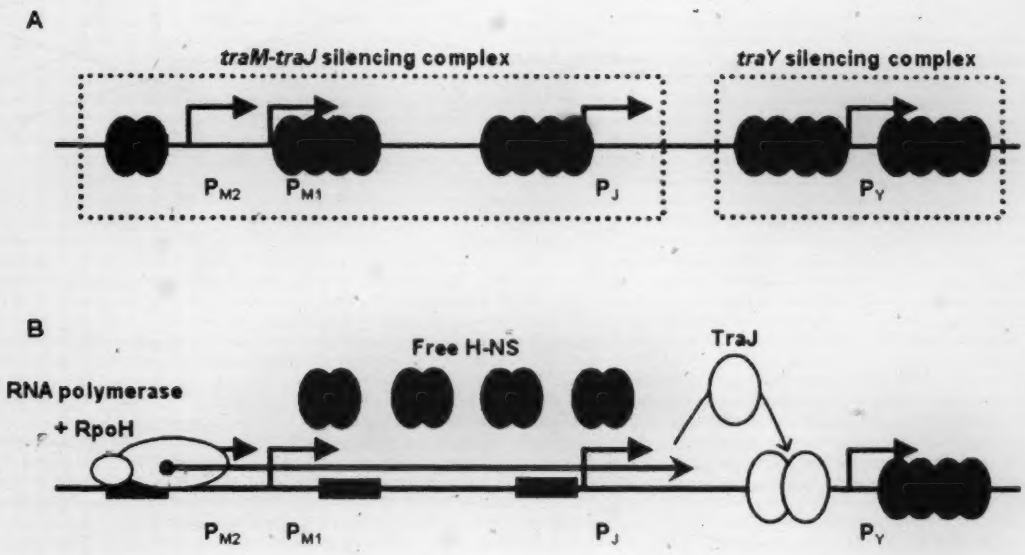
rapidly activated again once conditions improve. On the other hand, P_J upregulation could replenish the system of 'fresh' TraJ, while conditions improve. On the other hand, P_J upregulation could replenish the system of 'fresh' TraJ, while 'old' TraJ is eventually degraded. The loss of the C-terminus would lead to the smaller, inactive protein (22.49 kDa); however, mutants lacking the last 30 residues can still form dimers. TraJ seems to have multiple binding partners in an *hns* mutant or in the presence of F in the background. Whether these higher-order complexes involve other proteins or TraJ homo-oligomers is still unknown. The similarities between F TraJ and the MarR/SlyA family members, such as RovA, suggest TraJ belongs to this family of proteins. Both bind DNA and have an essential four amino acid tail at the C-terminus. In RovA, this tail is considered to be surface-exposed in order to interact, most probably, with other proteins (Tran *et al.*, 2005). Moreover, both can form dimers and have an HTH site, although the HTH site in RovA is part of a winged-helix domain (Ellison & Miller, 2006). Finally, both would be countersilencers of H-NS and could also have a role in the direct activation of the RNA polymerase, although this has not been demonstrated for TraJ. Thus, the hypothesis that TraJ possesses a winged-helix domain and is a member the family of RovA-like regulators requires further work.

RovA shares close structural similarity with regulators MarR and MexR, especially in the winged-helix domain (Tran *et al.*, 2005). However, unlike MarR, which is a repressor, RovA functions mainly as an activator by counter-silencing, as does its orthologue SlyA of *E. coli* and *Salmonella* (Fang & Rimsky, 2008). Moreover, it has been shown that SlyA requires another regulator in order to properly activate the *pagC* promoter of *Salmonella*, where it can bind DNA without alleviating H-NS binding. SlyA could modify the DNA topology just enough to allow the regulator PhoP to

enter the system and activate *pagC* transcription (Perez *et al.*, 2008). Thus, TraJ could also open the way for a second factor to join the complex and achieve transcription. This protein factor could also be part of the large complexes seen in crosslinking assays. As suggested before, this factor could be ArcA which is shown to have a binding site in the P_Y region (Strohmaier *et al.*, 1998), TraY, which regulates itself (Silverman & Sholl, 1996), or another nucleoid-associated protein like Lrp (Starcic-Erjavec *et al.*, 2003). Moreover, since silencing is closely related to nutritional stress, CRP could also be an important factor in this de-silencing complex as it has been shown to activate conjugative transfer in pRK100 (Starcic *et al.*, 2003).

H-NS silences F transfer when conditions are not favorable for the cell, as in stationary phase. How the transfer region is activated again once the stress is overcome is still under investigation. It is known that TraJ alleviates silencing of *traY*; however *traJ* is also silenced by H-NS in stationary phase (Will *et al.*, 2004). Moreover, the *traM* promoter is also part of this extensive repression with *traM* possessing two functional promoters (Penfold *et al.*, 1996). It has been shown that the lower the activity of the promoter, the stronger the repression by silencing, and vice versa (Nagarajavel *et al.*, 2007). Moreover, proper silencing requires a downstream regulatory element (DRE) and an upstream regulatory element (URE), also known as flanking nucleation sites (Nagarajavel *et al.*, 2007). Thus, since P_M is very strong (Lu *et al.*, 2003) as it is composed of two individual promoters (P_{M1} and P_{M2}) (Penfold *et al.*, 1996), it could overcome silencing more rapidly than the *traJ* or the *traY* promoters. This promoter could probably be activated with help from an alternative sigma factor, such as RpoH, which would help unwind the DNA thereby activating transcription of the *traM* gene by RNA polymerase (Figure 5.3). This

Figure 5.3 Model for silencing-desilencing of F transfer region. A) Silencing of the three main promoters (P_M, P_J and P_Y) by H-NS. Silencing of both the *traM* (1 and 2) promoters and the *traJ* promoter is coupled (shown by dotted rectangles). B) When silencing is countered, P_{M2}, having affinity for RpoH, overcomes first the H-NS nucleocomplex (clusters of dark circles). This disrupts the URE-DRE organization for silencing and allows RNA polymerase to read into *traJ* (long arrow), removing H-NS from P_J. TraJ is produced anew and alleviates H-NS silencing of P_Y, thus, activating conjugative transfer.


A
traM-traJ silencing complex
traY silencing complex
P_{M2}
P_{M1}
P_J
P_Y
B
Free H-NS
TraJ
RNA polymerase
+ RpoH
P_{M2}
P_{M1}
P_J
P_Y

alteration in superhelical density could also allow *traJ* to be de-silenced, probably by read-through transcripts from the *traM* promoter (Dempsey & Fee, 1990). Furthermore, these read-through transcripts from P_M into P_J suggest that a relationship exists between *traM* and *traJ* silencing (Will *et al.*, 2004). Both these promoters appear to be silenced within the same supercoiled repression complex. Thus, the loss of the P_M-H-NS complex that would function as the 'URE' would result in only the *traJ* DRE being bound by H-NS, which would function as a transcriptional roadblock. This kind of H-NS roadblock can be overcome, as the strength of the RNA polymerase when engaged in elongation is approximately twice as much as is needed to disrupt the H-NS nucleocomplex (Dame *et al.*, 2006). Finally, once *traJ* is free for transcription, TraJ would start de-silencing of the *traY* promoter. TraY would in turn increase P_M transcription which would give the burst of transcription needed to fully activate the *tra* control region.

B. Future directions

Future experiments would include protein-protein interactions between TraJ and itself, H-NS and core RNA polymerase. Thus, if TraJ binds the core RNA polymerase as a sigma factor, that would suggest another method of de-silencing. Furthermore, C-terminal mutants of TraJ could also be tested for binding with RNA polymerase in order to see if the C-terminal tail is directly interacting with the enzyme. Moreover, electrophoretic mobility shift assays (EMSA) with pure TraJ and different sizes of DNA fragments for TraJ binding could be tested, and the specific site of binding would be determined this way. Other proteins, like ArcA or TraY itself, could be tested for interaction with the TraJ-P_Y complex without H-NS being present. Finally, if TraJ was purified in enough quantity and in

a stable form, mass spectrometry could be tested again in order to determine the specific post-translational modification that inactivates TraJ in stationary phase.

Future experiments would also require the characterization of the C-terminal motif in TraJ. It might have an effect on the TraJ-RNA polymerase interactions; however this has not yet been proven. This domain was originally thought to be a dimerization domain but cross-linking experiments showed otherwise. A probable alternative for this dimerization domain could then be located near the N-terminus, as another cluster of conserved amino acids between different F-like TraJ is found there (residues 30-55). However, dimerization was maintained for all point mutants constructed, including P31A, which is inside this region (results not shown). If this domain proved not to be a dimerization domain, then further experiments would be needed to show its importance. Also, it would be of importance to determine the structure of TraJ by crystallization of the pure protein in order to know if the HTH site is part of a winged-helix domain, further supporting the similarities between TraJ and other members of the MarR/SlyA family of regulators.

TraJ cloned in the pET SUMO (Small Ubiquitin-like Modifier) plasmid (Invitrogen) has been recently constructed (D. Arutyunov, unpublished results). This SUMO-TraJ protein is soluble and is easily purified. Preliminary studies of the EMSA and cross-linking *in vitro* with the SUMO-TraJ system have been performed; however, proteolysis to remove the SUMO fusion renders the protein insoluble again. Thus, experiments with the SUMO-TraJ complex, where SUMO is fused to the N-terminus of TraJ, have been carried out instead and interesting results have been obtained. It would seem that SUMO-TraJ binds to DNA, albeit in a nonspecific

manner. SUMO-TraJ is also capable of forming dimers and higher order of oligomers, some unable to migrate in the gel. Moreover, the crystal structure of the SUMO-TraJ complex is also being pursued. In addition, the importance of the five cysteines for the formation of disulfide bonds is also being examined by mutagenesis in SUMO-TraJ. Further testing is needed in order to validate these results and to obtain the purified soluble TraJ protein.

As shown before, TraJ is digested by the HslVU protease, activated by CpxA, when the cell suffers extracytoplasmic stress (Lau-Wong *et al.*, 2008). Recent results have suggested that TraJ is stabilized and protected from HslVU digestion by the presence of F, with TraR being implicated (B. Beadle, personal communication). Moreover, it is also being tested whether TraR binds directly to TraJ using half-life experiments. Thus, it would be of interest to determine if the TraJ mutants created here have an effect on this interaction or if any of them are immune to HslVU digestion. Moreover, it will also be of interest to determine whether TraR, which seems to possess a zinc-finger DNA-binding motif, also has a role in countersilencing H-NS along TraJ.

As explained above, P_M could be de-silenced with the help of an alternative sigma factor. Moreover, it is not uncommon for dual promoters to have different affinities for different sigma factors (Foulger & Errington, 1991). Hence, there is a possibility that RpoH (σ^H) could assist P_M in alleviating repression. A base level of σ^H (approximately 20 molecules per cell) is present during normal growth (Straus *et al.*, 1987). This low level of σ^H could be enough to start the cascade of countersilencing events and restore conjugation under permissive conditions. Moreover, it has been shown that σ^H is required for *repE* transcription, which product is essential

for F replication (Wada *et al.*, 1987). The countersilencing effect of σ^H could also be an indirect one, as it could be required for the production of another important factor. This possibility is also being considered in current experiments.

My work here suggests a mechanism for TraJ function and has shown that TraJ belongs to the MarR/SlyA family because of the many characteristics they share. This is a good foundation for future experiments in the field of horizontal gene transfer and gene regulation.

Chapter 6. References

Achtman, M., Willetts, N. & Clark, A. J. (1971). Beginning a genetic analysis of conjugational transfer determined by the F factor in *Escherichia coli* by isolation and characterization of transfer-deficient mutants. *J Bacteriol* **106**, 529-538.

Achtman, M., Kennedy, N. & Skurray, R. (1977). Cell--cell interactions in conjugating *Escherichia coli*: role of TraT protein in surface exclusion. *Proc Natl Acad Sci U S A* **74**, 5104-5108.

Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S. & Ishihama, A. (1999). Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* **181**, 6361-6370.

Anthony, K. G., Klimke, W. A., Manchak, J. & Frost, L. S. (1999). Comparison of proteins involved in pilus synthesis and mating pair stabilization from the related plasmids F and R100-1: insights into the mechanism of conjugation. *J Bacteriol* **181**, 5149-5159.

Azam, T. A. & Ishihama, A. (1999). Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J Biol Chem* **274**, 33105-33113.

Blot, N., Mavathur, R., Geertz, M., Travers, A. & Muskhelishvili, G. (2006). Homeostatic regulation of supercoiling sensitivity coordinates transcription of the bacterial genome. *EMBO Rep* **7**, 710-715.

Bouffartigues, E., Buckle, M., Badaut, C., Travers, A. & Rimsky, S. (2007). H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nat Struct Mol Biol* **14**, 441-448.

Boyd, E. F., Hill, C. W., Rich, S. M. & Hartl, D. L. (1996). Mosaic structure of plasmids from natural populations of *Escherichia coli*. *Genetics* **143**, 1091-1100.

Bradley, D. E. (1980). Determination of pili by conjugative bacterial drug resistance plasmids of incompatibility groups B, C, H, J, K, M, V, and X. *J Bacteriol* **141**, 828-837.

Camacho, E. M. & Casadesus, J. (2002). Conjugal transfer of the virulence plasmid of *Salmonella enterica* is regulated by the leucine-responsive regulatory protein and DNA adenine methylation. *Mol Microbiol* **44**, 1589-1598.

Casadesus, J. & Low, D. (2006). Epigenetic gene regulation in the bacterial world. *Microbiol Mol Biol Rev* **70**, 830-856.

Chase, J. W., Merrill, B. M. & Williams, K. R. (1983). F sex factor encodes a single-stranded DNA binding protein (SSB) with extensive sequence homology to *Escherichia coli* SSB. *Proc Natl Acad Sci U S A* **80**, 5480-5484.

Cheah, K. C. & Skurray, R. (1986). The F plasmid carries an IS3 insertion within finO. *J Gen Microbiol* **132**, 3269-3275.

Cuozzo, M., Silverman, P. M. & Minkley, E. G., Jr. (1984). Overproduction in *Escherichia coli* K-12 and purification of the TraJ protein encoded by the conjugative plasmid F. *J Biol Chem* **259**, 6659-6666.

Cuozzo, M. & Silverman, P. M. (1986). Characterization of the F plasmid TraJ protein synthesized in F' and Hfr strains of *Escherichia coli* K-12. *J Biol Chem* **261**, 5175-5179.

Dame, R. T. & Goosen, N. (2002). HU: promoting or counteracting DNA compaction? *FEBS Lett* **529**, 151-156.

Dame, R. T., Noom, M. C. & Wuite, G. J. (2006). Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature* **444**, 387-390.

Daubin, V., Lerat, E. & Perriere, G. (2003). The source of laterally transferred genes in bacterial genomes. *Genome Biol* **4**, R57.

Dempsey, W. B. & Fee, B. E. (1990). Integration host factor affects expression of two genes at the conjugal transfer origin of plasmid R100. *Mol Microbiol* **4**, 1019-1028.

Dempsey, W. B. (1994). *traJ* sense RNA initiates at two different promoters in R100-1 and forms two stable hybrids with antisense *finP* RNA. *Mol Microbiol* **13**, 313-326.

Dersch, P., Schmidt, K. & Bremer, E. (1993). Synthesis of the *Escherichia coli* K-12 nucleoid-associated DNA-binding protein H-NS is subjected to growth-phase control and autoregulation. *Mol Microbiol* **8**, 875-889.

Dorman, C. J., Hinton, J. C. & Free, A. (1999). Domain organization and oligomerization among H-NS-like nucleoid-associated proteins in bacteria. *Trends Microbiol* **7**, 124-128.

Dorman, C. J. (2004). H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol* **2**, 391-400.

Dutreix, M., Backman, A., Celerier, J., Bagdasarian, M. M., Sommer, S., Bailone, A., Devoret, R. & Bagdasarian, M. (1988). Identification of *psiB* genes of plasmids F and R6-5. Molecular basis for *psiB* enhanced expression in plasmid R6-5. *Nucleic Acids Res* **16**, 10669-10679.

Ellison, D. W. & Miller, V. L. (2006). Regulation of virulence by members of the MarR/SlyA family. *Curr Opin Microbiol* **9**, 153-159.

Fang, F. C. & Rimsky, S. (2008). New insights into transcriptional regulation by H-NS. *Curr Opin Microbiol* **11**, 113-120.

Finlay, B. B., Frost, L. S., Paranchych, W. & Willetts, N. S. (1986). Nucleotide sequences of five IncF plasmid *finP* alleles. *J Bacteriol* **167**, 754-757.

Firth, N. & Skurray, R. (1992). Characterization of the F plasmid bifunctional conjugation gene, traG. *Mol Gen Genet* **232**, 145-153.

Foulger, D. & Errington, J. (1991). Sequential activation of dual promoters by different sigma factors maintains *spoVJ* expression during successive developmental stages of *Bacillus subtilis*. *Mol Microbiol* **5**, 1363-1373.

Frost, L. S., Ippen-Ihler, K. & Skurray, R. A. (1994). Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol Rev* **58**, 162-210.

Frost, L. S. & Manchak, J. (1998). F- phenocopies: characterization of expression of the F transfer region in stationary phase. *Microbiology* **144 (Pt 9)**, 2579-2587.

Gajiwala, K. S. & Burley, S. K. (2000). Winged helix proteins. *Curr Opin Struct Biol* **10**, 110-116.

Ghetu, A. F., Gubbins, M. J., Oikawa, K., Kay, C. M., Frost, L. S. & Glover, J. N. (1999). The FinO repressor of bacterial conjugation contains two RNA binding regions. *Biochemistry* **38**, 14036-14044.

Ghetu, A. F., Gubbins, M. J., Frost, L. S. & Glover, J. N. (2000). Crystal structure of the bacterial conjugation repressor FinO. *Nat Struct Biol* **7**, 565-569.

Ghetu, A. F., Arthur, D. C., Kerppola, T. K. & Glover, J. N. (2002). Probing FinO-FinP RNA interactions by site-directed protein-RNA crosslinking and gelFRET. *RNA* **8**, 816-823.

Gubbins, M. J., Lau, I., Will, W. R., Manchak, J. M., Raivio, T. L. & Frost, L. S. (2002). The positive regulator, TraJ, of the *Escherichia coli* F plasmid is unstable in a *cpxA** background. *J Bacteriol* **184**, 5781-5788.

Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**, 4121-4130.

Ham, L. M., Cram, D. & Skurray, R. (1989). Transcriptional analysis of the F plasmid surface exclusion region: mapping of *traS, traT,* and *traD* transcripts. *Plasmid* **21**, 1-8.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557-580.

Heroven, A. K., Nagel, G., Tran, H. J., Parr, S. & Dersch, P. (2004). RovA is autoregulated and antagonizes H-NS-mediated silencing of invasin and *rovA* expression in *Yersinia pseudotuberculosis*. *Mol Microbiol* **53**, 871-888.

Hommais, F., Krin, E., Laurent-Winter, C., Soutourina, O., Malpertuy, A., Le Caer, J. P., Danchin, A. & Bertin, P. (2001). Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol Microbiol* **40**, 20-36.

Howard, M. T., Nelson, W. C. & Matson, S. W. (1995). Stepwise assembly of a relaxosome at the F plasmid origin of transfer. *J Biol Chem* **270**, 28381-28386.

Inamoto, S., Yoshioka, Y. & Ohtsubo, E. (1988). Identification and characterization of the products from the *traJ* and *traY* genes of plasmid R100. *J Bacteriol* **170**, 2749-2757.

Jalajakumari, M. B., Guidolin, A., Buhk, H. J., Manning, P. A., Ham, L. M., Hodgson, A. L., Cheah, K. C. & Skurray, R. A. (1987). Surface exclusion genes *traS* and *traT* of the F sex factor of *Escherichia coli* K-12. Determination of the nucleotide sequence and promoter and terminator activities. *J Mol Biol* **198**, 1-11.

Jerome, L. J. & Frost, L. S. (1999). In vitro analysis of the interaction between the FinO protein and FinP antisense RNA of F-like conjugative plasmids. *J Biol Chem* **274**, 10356-10362.

Jerome, L. J., van Biesen, T. & Frost, L. S. (1999). Degradation of FinP antisense RNA from F-like plasmids: the RNA-binding protein, FinO, protects FinP from ribonuclease E. *J Mol Biol* **285**, 1457-1473.

Johansson, J. & Uhlin, B. E. (1999). Differential protease-mediated turnover of H-NS and StpA revealed by a mutation altering protein stability and stationary-phase survival of *Escherichia coli*. *Proc Natl Acad Sci U S A* **96**, 10776-10781.

Klimke, W. A. & Frost, L. S. (1998). Genetic analysis of the role of the transfer gene, *traN*, of the F and R100-1 plasmids in mating pair stabilization during conjugation. *J Bacteriol* **180**, 4036-4043.

Klimke, W. A., Rypien, C. D., Klinger, B., Kennedy, R. A., Rodriguez-Maillard, J. M. & Frost, L. S. (2005). The mating pair stabilization protein, TraN, of the F plasmid is an outer-membrane protein with two regions that are important for its function in conjugation. *Microbiology* **151**, 3527-3540.

Koraimann, G., Koraimann, C., Koronakis, V., Schlager, S. & Hogenauer, G. (1991). Repression and derepression of conjugation of plasmid R1 by wild-type and mutated *finP* antisense RNA. *Mol Microbiol* **5**, 77-87.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lahue, E. E. & Matson, S. W. (1990). Purified *Escherichia coli* F-factor TraY protein binds *oriT*. *J Bacteriol* **172**, 1385-1391.

Lai, E., Prezioso, V. R., Smith, E., Litvin, O., Costa, R. H. & Darnell, J. E., Jr. (1990). HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally. *Genes Dev* **4**, 1427-1436.

Lai, E., Clark, K. L., Burley, S. K. & Darnell, J. E., Jr. (1993). Hepatocyte nuclear factor 3/fork head or "winged helix" proteins: a family of transcription factors of diverse biologic function. *Proc Natl Acad Sci U S A* **90**, 10421-10423.

Lang, B., Blot, N., Bouffartigues, E. & other authors (2007). High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Res* **35**, 6330-6337.

Lau-Wong, I. C., Locke, T., Ellison, M. J., Raivio, T. L. & Frost, L. S. (2008). Activation of the Cpx regulon destabilizes the F plasmid transfer activator, TraJ, via the HslVU protease in *Escherichia coli*. *Mol Microbiol* **67**, 516-527.

Laurenson, P. & Rine, J. (1992). Silencers, silencing, and heritable transcriptional states. *Microbiol Rev* **56**, 543-560.

Lawley, T. D., Klimke, W. A., Gubbins, M. J. & Frost, L. S. (2003). F factor conjugation is a true type IV secretion system. *FEMS Microbiol Lett* **224**, 1-15.

Loh, S., Skurray, R., Celerier, J., Bagdasarian, M., Bailone, A. & Devoret, R. (1990). Nucleotide sequence of the *psiA* (plasmid SOS inhibition) gene located on the leading region of plasmids F and R6-5. *Nucleic Acids Res* **18**, 4597.

Low, D. A., Weyand, N. J. & Mahan, M. J. (2001). Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. *Infect Immun* **69**, 7197-7204.

Lu, J., Fekete, R. A. & Frost, L. S. (2003). A rapid screen for functional mutants of TraM, an autoregulatory protein required for F conjugation. *Mol Genet Genomics* **269**, 227-233.

Lu, J., Edwards, R. A., Wong, J. J., Manchak, J., Scott, P. G., Frost, L. S. & Glover, J. N. (2006). Protonation-mediated structural flexibility in the F conjugation regulatory protein, TraM. *EMBO J* **25**, 2930-2939.

Lucchini, S., Rowley, G., Goldberg, M. D., Hurd, D., Harrison, M. & Hinton, J. C. (2006). H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog* **2**, e81.

Lum, P. L. & Schildbach, J. F. (1999). Specific DNA recognition by F Factor TraY involves beta-sheet residues. *J Biol Chem* **274**, 19644-19648.

Lum, P. L., Rodgers, M. E. & Schildbach, J. F. (2002). TraY DNA recognition of its two F factor binding sites. *J Mol Biol* **321**, 563-578.

Majdalani, N. & Ippen-Ihler, K. (1996). Membrane insertion of the F-pilin subunit is Sec independent but requires leader peptidase B and the proton motive force. *J Bacteriol* **178**, 3742-3747.

Manwaring, N. P., Skurray, R. A. & Firth, N. (1999). Nucleotide sequence of the F plasmid leading region. *Plasmid* **41**, 219-225.

Martinez-Hackert, E. & Stock, A. M. (1997). Structural relationships in the OmpR family of winged-helix transcription factors. *J Mol Biol* **269**, 301-312.

Masai, H. & Arai, K. (1997). F*rpo*: a novel single-stranded DNA promoter for transcription and for primer RNA synthesis of DNA replication. *Cell* **89**, 897-907.

Matson, S. W. & Ragonese, H. (2005). The F-plasmid TraI protein contains three functional domains required for conjugative DNA strand transfer. *J Bacteriol* **187**, 697-706.

Maurelli, A. T. & Sansonetti, P. J. (1988). Identification of a chromosomal gene controlling temperature-regulated expression of *Shigella* virulence. *Proc Natl Acad Sci U S A* **85**, 2820-2824.

Meynell, E., Meynell, G. G. & Datta, N. (1968). Phylogenetic relationships of drug-resistance factors and other transmissible bacterial plasmids. *Bacteriol Rev* **32**, 55-83.

Moller, T., Franch, T., Hojrup, P., Keene, D. R., Bachinger, H. P., Brennan, R. G. & Valentin-Hansen, P. (2002). Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol Cell* **9**, 23-30.

Moore, D., Maneewannakul, K., Maneewannakul, S., Wu, J. H., Ippen-Ihler, K. & Bradley, D. E. (1990). Characterization of the F-plasmid conjugative transfer gene *traU*. *J Bacteriol* **172**, 4263-4270.

Muffler, A., Fischer, D. & Hengge-Aronis, R. (1996). The RNA-binding protein HF-I, known as a host factor for phage Qbeta RNA replication, is essential for *rpoS* translation in *Escherichia coli*. *Genes Dev* **10**, 1143-1151.

Nagarajavel, V., Madhusudan, S., Dole, S., Rahmouni, A. R. & Schnetz, K. (2007). Repression by binding of H-NS within the transcription unit. *J Biol Chem* **282**, 23622-23630.

Navarre, W. W., Porwollik, S., Wang, Y., McClelland, M., Rosen, H., Libby, S. J. & Fang, F. C. (2006). Selective silencing of foreign DNA with low GC content by the H-NS protein in Salmonella. *Science* **313**, 236-238.

Navarre, W. W., McClelland, M., Libby, S. J. & Fang, F. C. (2007). Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes Dev* **21**, 1456-1471.

Nelson, W. C., Morton, B. S., Lahue, E. E. & Matson, S. W. (1993). Characterization of the *Escherichia coli* F factor *traY* gene product and its binding sites. *J Bacteriol* **175**, 2221-2228.

Nelson, W. C. & Matson, S. W. (1996). The F plasmid *traY* gene product binds DNA as a monomer or a dimer: structural and functional implications. *Mol Microbiol* **20**, 1179-1187.

Nieto, J. M., Mourino, M., Balsalobre, C., Madrid, C., Prenafeta, A., Munoa, F. J. & Juarez, A. (1997). Construction of a double *hha hns* mutant of *Escherichia coli*: effect on DNA supercoiling and alpha-haemolysin production. *FEMS Microbiol Lett* **155**, 39-44.

Nieto, J. M., Madrid, C., Miquelay, E., Parra, J. L., Rodriguez, S. & Juarez, A. (2002). Evidence for direct protein-protein interaction between members of the enterobacterial Hha/YmoA and H-NS families of proteins. *J Bacteriol* **184**, 629-635.

Niu, W., Kim, Y., Tau, G., Heyduk, T. & Ebright, R. H. (1996). Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase. *Cell* **87**, 1123-1134.

Ohlendorf, D. H., Anderson, W. F. & Matthews, B. W. (1983). Many gene-regulatory proteins appear to have a similar alpha-helical fold that

binds DNA and evolved from a common precursor. *J Mol Evol* **19**, 109-114.

Pabo, C. O. & Sauer, R. T. (1984). Protein-DNA recognition. *Annu Rev Biochem* **53**, 293-321.

Paytubi, S., Madrid, C., Forns, N., Nieto, J. M., Balsalobre, C., Uhlin, B. E. & Juarez, A. (2004). YdgT, the Hha paralogue in *Escherichia coli*, forms heteromeric complexes with H-NS and StpA. *Mol Microbiol* **54**, 251-263.

Penfold, S. S., Simon, J. & Frost, L. S. (1996). Regulation of the expression of the *traM* gene of the F sex factor of *Escherichia coli*. *Mol Microbiol* **20**, 549-558.

Perez, J. C., Latifi, T. & Groisman, E. A. (2008). Overcoming H-NS-mediated transcriptional silencing of horizontally acquired genes by the PhoP and SlyA proteins in *Salmonella enterica*. *J Biol Chem* **283**, 10773-10783.

Prosseda, G., Falconi, M., Giangrossi, M., Gualerzi, C. O., Micheli, G. & Colonna, B. (2004). The *virF* promoter in *Shigella*: more than just a curved DNA stretch. *Mol Microbiol* **51**, 523-537.

Pul, U., Wurm, R., Lux, B., Meltzer, M., Menzel, A. & Wagner, R. (2005). LRP and H-NS--cooperative partners for transcription regulation at *Escherichia coli* rRNA promoters. *Mol Microbiol* **58**, 864-876.

Raibaud, O. & Schwartz, M. (1984). Positive control of transcription initiation in bacteria. *Annu Rev Genet* **18**, 173-206.

Raivio, T. L. & Silhavy, T. J. (1997). Transduction of envelope stress in *Escherichia coli* by the Cpx two-component system. *J Bacteriol* **179**, 7724-7733.

Reeves, P. & Willetts, N. (1974). Plasmid specificity of the origin of transfer of sex factor F. *J Bacteriol* **120**, 125-130.

Sandercock, J. R. & Frost, L. S. (1998). Analysis of the major domains of the F fertility inhibition protein, FinO. *Mol Gen Genet* **259**, 622-629.

Shin, M., Song, M., Rhee, J. H., Hong, Y., Kim, Y. J., Seok, Y. J., Ha, K. S., Jung, S. H. & Choy, H. E. (2005). DNA looping-mediated repression

by histone-like protein H-NS: specific requirement of Esigma70 as a cofactor for looping. *Genes Dev* **19**, 2388-2398.

Silverman, P. M., Wickersham, E. & Harris, R. (1991). Regulation of the F plasmid *traY* promoter in *Escherichia coli* by host and plasmid factors. *J Mol Biol* **218**, 119-128.

Silverman, P. M. & Sholl, A. (1996). Effect of *traY* amber mutations on F-plasmid *traY* promoter activity in vivo. *J Bacteriol* **178**, 5787-5789.

Somers, W. S. & Phillips, S. E. (1992). Crystal structure of the met repressor-operator complex at 2.8 A resolution reveals DNA recognition by beta-strands. *Nature* **359**, 387-393.

Starcic-Erjavec, M., van Putten, J. P., Gaastra, W., Jordi, B. J., Grabnar, M. & Zgur-Bertok, D. (2003). H-NS and Lrp serve as positive modulators of *traJ* expression from the *Escherichia coli* plasmid pRK100. *Mol Genet Genomics* **270**, 94-102.

Starcic, M., Zgur-Bertok, D., Jordi, B. J., Wosten, M. M., Gaastra, W. & van Putten, J. P. (2003). The cyclic AMP-cyclic AMP receptor protein complex regulates activity of the *traJ* promoter of the *Escherichia coli* conjugative plasmid pRK100. *J Bacteriol* **185**, 1616-1623.

Steitz, T. A., Ohlendorf, D. H., McKay, D. B., Anderson, W. F. & Matthews, B. W. (1982). Structural similarity in the DNA-binding domains of catabolite gene activator and *cro* repressor proteins. *Proc Natl Acad Sci U S A* **79**, 3097-3100.

Strahl-Bolsinger, S., Hecht, A., Luo, K. & Grunstein, M. (1997). SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. *Genes Dev* **11**, 83-93.

Straus, D. B., Walter, W. A. & Gross, C. A. (1987). The heat shock response of *E. coli* is regulated by changes in the concentration of sigma 32. *Nature* **329**, 348-351.

Strohmaier, H., Noiges, R., Kotschan, S., Sawers, G., Hogenauer, G., Zechner, E. L. & Koraimann, G. (1998). Signal transduction and bacterial conjugation: characterization of the role of ArcA in regulating conjugative transfer of the resistance plasmid R1. *J Mol Biol* **277**, 309-316.

Taki, K., Abo, T. & Ohtsubo, E. (1998). Regulatory mechanisms in expression of the *traY-I* operon of sex factor plasmid R100: involvement of *traJ* and *traY* gene products. *Genes Cells* **3**, 331-345.

Tani, T. H., Khodursky, A., Blumenthal, R. M., Brown, P. O. & Matthews, R. G. (2002). Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. *Proc Natl Acad Sci U S A* **99**, 13471-13476.

Tatum, E. L. & Lederberg, J. (1947). Gene Recombination in the Bacterium *Escherichia coli*. *J Bacteriol* **53**, 673-684.

Tendeng, C. & Bertin, P. N. (2003). H-NS in Gram-negative bacteria: a family of multifaceted proteins. *Trends Microbiol* **11**, 511-518.

Tobe, T., Yoshikawa, M., Mizuno, T. & Sasakawa, C. (1993). Transcriptional control of the invasion regulatory gene *virB* of *Shigella flexneri*: activation by *virF* and repression by H-NS. *J Bacteriol* **175**, 6142-6149.

Torreblanca, J., Marques, S. & Casadesus, J. (1999). Synthesis of FinP RNA by plasmids F and pSLT is regulated by DNA adenine methylation. *Genetics* **152**, 31-45.

Towbin, H., Staehelin, T. & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350-4354.

Tran, H. J., Heroven, A. K., Winkler, L., Spreter, T., Beatrix, B. & Dersch, P. (2005). Analysis of RovA, a transcriptional regulator of *Yersinia pseudotuberculosis* virulence that acts through antirepression and direct transcriptional activation. *J Biol Chem* **280**, 42423-42432.

van Biesen, T. & Frost, L. S. (1994). The FinO protein of IncF plasmids binds FinP antisense RNA and its target, *traJ* mRNA, and promotes duplex formation. *Mol Microbiol* **14**, 427-436.

Wada, C., Imai, M. & Yura, T. (1987). Host control of plasmid replication: requirement for the sigma factor sigma 32 in transcription of mini-F replication initiator gene. *Proc Natl Acad Sci U S A* **84**, 8849-8853.

Watanabe, T., Nishida, H., Ogata, C., Arai, T. & Sato, S. (1964). Episome-Mediated Transfer of Drug Resistance in Enterobacteriaceae. Vii. Two Types of Naturally Occurring R Factors. *J Bacteriol* **88**, 716-726.

Will, W. R., Lu, J. & Frost, L. S. (2004). The role of H-NS in silencing F transfer gene expression during entry into stationary phase. *Mol Microbiol* **54**, 769-782.

Will, W. R. & Frost, L. S. (2006a). Characterization of the opposing roles of H-NS and TraJ in transcriptional regulation of the F-plasmid *tra* operon. *J Bacteriol* **188**, 507-514.

Will, W. R. & Frost, L. S. (2006b). Hfq is a regulator of F-plasmid TraJ and TraM synthesis in *Escherichia coli*. *J Bacteriol* **188**, 124-131.

Willetts, N. & Maule, J. (1986). Specificities of IncF plasmid conjugation genes. *Genet Res* **47**, 1-11.

Williams, R. M., Rimsky, S. & Buc, H. (1996). Probing the structure, function, and interactions of the *Escherichia coli* H-NS and StpA proteins by using dominant negative derivatives. *J Bacteriol* **178**, 4335-4343.

Williams, R. M. & Rimsky, S. (1997). Molecular aspects of the *E. coli* nucleoid protein, H-NS: a central controller of gene regulatory networks. *FEMS Microbiol Lett* **156**, 175-185.

Williamson, H. S. & Free, A. (2005). A truncated H-NS-like protein from enteropathogenic *Escherichia coli* acts as an H-NS antagonist. *Mol Microbiol* **55**, 808-827.

Wu, J. H. & Ippen-Ihler, K. (1989). Nucleotide sequence of *traQ* and adjacent loci in the *Escherichia coli* K-12 F-plasmid transfer operon. *J Bacteriol* **171**, 213-221.

Zheng, D., Constantinidou, C., Hobman, J. L. & Minchin, S. D. (2004). Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Res* **32**, 5874-5893.